CONFORMATIONAL SEARCH OF PROTEINS AND PROTEIN LOOPS

By

Ranjitha Venkataramani

Submitted to the Department of Chemistry and the Faculty of

the Graduate School of the University of Kansas

in partial fulfillment of the requirements for the degree of

Master of Science

_____
(Chair)

_____

_____
(Committee Members)

Date Submitted: _____

ACKNOWLEDGEMENTS

I would like to profusely thank my advisor Dr. Krzysztof Kuczera for his continued support and guidance. I would also like to express my gratitude to Dr. Mario Rivera and Dr. David R. Benson for their valuable suggestions. I would like to thank Yao Athanase Houndonougbo for his encouragement and support. Finally I would like to thank my friends Krishna Priya Kotcherlakota, Karthikeyan Ganapathy and Kuyil Mozhi Nangai.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS……..*2*

ABSTRACT…….*6*

1. INTRODUCTION……...……..8

1.1 Cytochrome b5…….9

1.2 Microsomal and Outer Mitochondrial cytochrome b5……...10

1.2.1 Structures .....…….10

1.2.2 Biophysical properties…….……..…..12

1.3 Motivation……..13

1.4 Molecular dynamics simulations…….13

1.5 Replica exchange molecular dynamics……...15

1.6 Apocytochrome b5 dynamics…….….16

1.7 Holo protein dynamics..... 18

2. SIMULATION PROCEDURE 20

2.1 MD Simulations…… ……...20

2.1.1 Initialization……...21

2.1.2 Addition of ions 21

2.1.3 Solvating the protein..... 22

2.1.4 Equilibration of system..... 22

2.1.5 Molecular dynamics simulations……...…...……...23

2.2 REPLICA EXCHANGE MOLECULAR DYNAMICS……..24

2.1.1 Background ……....24

3. ANALYSIS OF PROTEIN DYNAMICS……...…….28

3.1 Protein structure analysis …….28

3.2 Analysis of core 1 and core 2……..……....31

3.3 RMS Fluctuations……..…….……..……...33

3.4 Radius of gyration… …….……39

3.5 Secondary structure analysis…….…….……41

3.6 Solvent accessibility...…….……46

3.7 Interresidue distance……..……...……...50

3.8 Trajectory of 20 NMR residues ……..……...……....51

3.9 Discussion …….……..……...……...….54

3.10 Conclusions……..……..……..……...…...58

3.11 Future work……..……..……....……59

4. REMD SIMULATIONS OF HOLO PROTEINS …..……..……..60

4.1 RMSF of cytochrome b5…….60

4.2 RMSD of backbone atoms …….63

4.3 Radius of gyration……..……...…...……64

4.4 RMSD histograms.……..……66

4.5 Energy histograms……. 68

4.6 Analysis of core 1 and core2 ……..……70

4.7 SASA of entire protein…..……..…….…....76

4.8 Native contacts…….……...……79

4.9 Secondary structure analysis……82

4.10 Comparison of Apo and Holo protein dynamics…….……..87

4.11 Discussion of results…….……...…….91

4.12 Conclusions…….……...93

5. MODELING PROTEIN LOOPS ……..……...……..96

5.1 Method……..……..…….……96

5.2 Results……...…..99

5.3 Conclusion & future work…….……...…102

REFERENCES……...103

ABSTRACT

This thesis focuses on understanding the structure and dynamics of proteins using simulations. The main system of interest in our case was cytochrome b_5, an electron transport protein, which exists in both Microsomal and Outer Mitochondrial isoforms. We conducted 20 ns Molecular Dynamics simulations of apocytochrome b_5 (protein without the heme group). The apocytochrome b_5 proteins were more flexible than the holo proteins and the effect of heme removal was mostly confined to the heme-binding (core 1) region. Increased mobility was also observed in the β5 sheet region residues (51-55) of the Microsomal apoprotein, whereas in the apo OM protein this beta sheet region was quite rigid. Consistent with experimental results, the heme-binding pocket in OM apoprotein was found to be more compact as reflected by lower radius of gyration and distances between heme ligands.

The second part of my thesis consisted of performing replica exchange simulations of the holo proteins. About 4.4 ns simulations of the proteins were conducted with 12 replicas spanning 320 K-450 K using the GB/SA implicit solvent model. The rat Microsomal protein exhibited larger RMS Deviations from the starting structure as compared to the rat OM protein. The variation of structure with respect to temperature showed that the rat OM protein had higher melting temperatures than rat Mc protein and the rat Microsomal protein was more susceptible to thermal unfolding.

The thermal unfolding process was observed to occur in a cooperative manner, wherein changes in the heme-binding environment also resulted in fluctuations in core 2. The last part of my thesis was reconstruction of selected protein loops using Replica Exchange simulations. The Effective Energy Function (EEF1) implicit solvation model was used to conduct these simulations. Conformational analysis of the simulated proteins was conducted and various conformations were ranked based on lowest energy conformer and lowest RMSD value. Several conformations, which were comparable with experimental structures, were generated. However, we couldn't arrive at best possible conformation due to lack of good scoring function.

1. INTRODUCTION

The different spatial arrangements of proteins due to free rotations around a single chemical bond are called conformations. Conformational search refers to the process of identifying all the lowest energy conformations from a set of possible conformations.[1] The global minimum structure is the overall lowest energy structure with reference to the energy domain of the protein. The multidimensional energy surface of proteins consists of several minima and energy barriers, hence arriving at the global minimum conformation is quite difficult.[2] Along with finding the global minimum structure, it is also important to find the local minimum structures that are likely to be populated under physiological conditions, as several properties can be derived using the ensemble average (statistical average of all the conformations).

Conformational search methods can be broadly classified into stochastic and deterministic, of which techniques like Monte Carlo simulations, simulated annealing, stochastic dynamics come under the stochastic category while Systematic search and molecular dynamics come under the deterministic category.[1] The first part of my thesis concerns molecular dynamics simulations of the cytochrome b_5 proteins. MD simulations were conducted on both the apo and holo forms of proteins. The main aim of this project was to compare the dynamics of the Microsomal and Outer Mitochondrial isoforms of rat proteins.

The second part of my thesis consists of modeling protein loops using Replica Exchange Molecular Dynamics and conformational analysis of the results. The aim of this project was to reconstruct protein loop regions using constrained minimization and using scoring methods arrive at best possible conformation.

1.1 Cytochrome b_5

Cytochromes b_5 are electron transport proteins, which contain a prosthetic heme group. [3,4] Though initially found in cecropia silkworm, it was named as cytochrome m, as it was excessively found in liver microsomes.[5,6] The cytochrome b5 proteins found in erythrocytes and bacteria are water-soluble whereas they exist in membrane bound forms in Outer Mitochondria and Endoplasmic Reticulum.[7]

The electron transport process refers to the successive migration of electrons from one protein to another by a series of oxidation-reduction reactions.[8] The Fe metal ion present in the heme group can interconvert between oxidation states (Fe^{2+} and Fe^{3+}), thereby enabling cytochrome b5 to function as an electron transport protein. Cytochromes b5 are involved in several biochemical reactions including anabolic metabolism of fats, steroids, P450 dependent reactions, biosynthesis of plasmalogen, cholesterol etc.[9,10]

1.2 Microsomal and Outer Mitochondrial cytochrome b5

From earlier studies, it is known that two different genes code for the isoforms of cytochrome b5.[11,12] The first gene codes for the Microsomal cytochrome b5, which is tail, anchored to the membrane of Endoplasmic reticulum, whereas the second gene specifically codes for the Outer Mitochondrial membrane cytochrome b5.[11,12] Both these isoforms are described by an N-terminal domain, which extends into the cytosol, a stretch of hydrophobic residues, which are embedded into the membrane acting as an anchor and finally a short C – terminal segment consisting of hydrophilic residues.

From mutation studies it has been found that the C - terminal residues determines whether the protein is localized within the ER and the presence of a positively charged residue resulted in protein being localized to outer membrane of mitochondria.[13] Through recent genome sequencing projects Microsomal and OM cytochrome b5 genes have been identified in wide range of phyla including plants, mammals, insects, birds, fish, and amphibians. [14]

1.2.1 Structures

The X-ray crystal structures of Microsomal cytochrome b5 and rat OM cytochrome b5 are highly similar with a backbone atoms RMS deviation of about 0.6 Angstroms. The conserved fold in cytochrome b5 belongs to the $\alpha + \beta$ class, and consists of two major hydrophobic cores which are separated by an antiparallel β sheet. The larger

hydrophobic core, also called as core 1, consists of the heme-binding pocket surrounded by four helices α2 – α5, and the five-stranded beta sheet β5. The fifth, sixth ligands of the heme iron are two histidines His 39 and His 63.[16]

The smaller hydrophobic core comprising of α1 and α6 and beta sheets β1 – β4, is believed to be responsible for the structural integrity of the protein. Sequence similarity between the mammalian cytochrome b5 (bovine Mc and rat OM) using BLASTP revealed that the sequences were 70% similar.[17] Also the side chain torsional angles are very similar, differing less than 60^{0}.[18]

The structures of bovine Microsomal cytochrome b5 and Outer Mitochondrial cytochrome b5 are shown below

Fig 1 a: Bovine microsomal cytochrome b5 – PDB: 1CYO

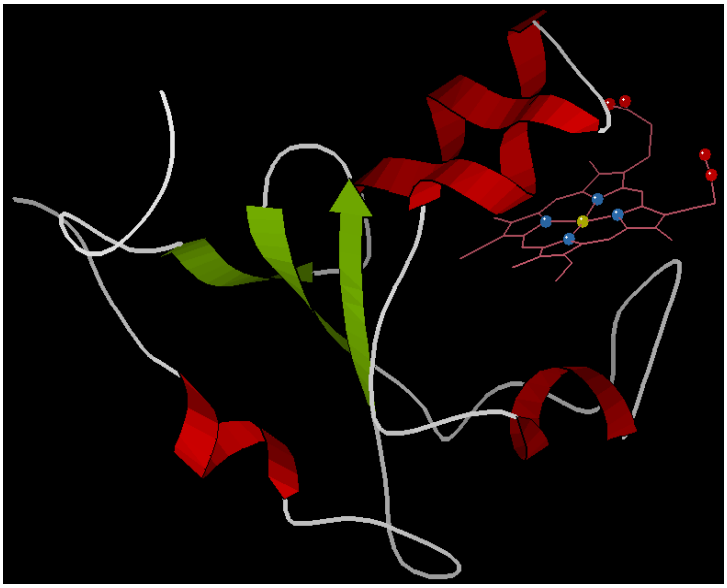


Fig 1 b: Rat Outer Mitochondrial cytochrome b5 – PDB: 1B5M

1.2.2 Biophysical properties

Though the 3D structures of microsomal cytochrome b5 and outer mitochondrial protein are very similar, considerable differences in physical properties have been observed. The rat outer mitochondrial cytochrome b5 protein has about 100mV more negative reduction potential as compared to microsomal cytochrome b5.[19] Also when subjected to thermal and chemical stability tests, it was observed that the rat OM protein is more stable than the Microsomal form.[20] Another major variation is the faster rate of hemin release in case of Microsomal cytochrome b5, whereas transfer of hemin from OM cytochrome b5 to apomyoglobin was unobservable at pH 7.0 and was very slow at pH 5.0. [20,21]

Apocytochrome b5 studies revealed that both the isoforms were equally stable at 25 ^{o} C in the absence of the heme group.[21] Hence the heme interaction with the protein seems to play an important role in the determining the stability of protein. The differences in biophysical property may explain the functionally divergent behavior of these two forms.

1.3 Motivation

To further understand these differences, we decided to conduct MD simulations of both holo cytochrome b5 and apocytochrome b5 (protein without the heme group). The aim of this project were as follows

1. To understand the differences between isoforms, which contribute towards the difference in biophysical properties like stability and hemin release.
2. To understand the effect of heme interaction with the cytochrome b5 .
3. To better understand the dynamics and internal motions of the proteins.
4. To study the effect of temperature on the dynamics and stability of proteins.

1.4 Molecular Dynamics Simulations

Molecular Dynamics (MD) simulations provide detailed information about atomic fluctuations, side chain fluctuations, domain motions, loop motions and conformational changes of proteins and nucleic acids. Advancements in parallel computing have made it possible to conduct MD simulations on complex bio-molecules. In case of molecular dynamics simulations we start with a crystal or NMR

structure and at regular intervals the classical equations of motion are solved and new velocities and position coordinates are calculated. This process is iteratively performed for a number of steps, thereby providing us a trajectory of the dynamic behavior of the protein. The rat OM protein consists of about 146 amino acid residues and is longer than rat Mc protein that has about 134 residues. The soluble part of the protein is about 90 residues in case of the Microsomal cytochrome b5 and 86 residues in case of OM cytochrome b5. The presence of NMR structure for Microsomal cytochrome b5 and X-ray crystal structure for the Outer Mitochondrial membrane protein have proved as a starting point to conduct MD simulations on these protein systems.

The MD simulations of cytochrome b5 proteins have been conducted using the CHARMM force fields and in explicit TIP3P solvent model to consider the solvent effects. Explicit counter ions are added to neutralize the system, thereby mimicking the biological environment. Previously several Molecular simulations have been conducted on cytochrome b5. Dagett and Storch performed 2.5 ns simulations of bovine Mc cytochrome b5. During their simulations, they discovered opening of a cleft, which exposed the internal hydrophobic residues and part of the heme.[16] Cheng et al also observed similar results during their 8.5 ns molecular dynamics simulation of bovine Microsomal cytochrome b5.[23]

In order to understand the folding process and the consequence of heme removal, molecular dynamics was conducted on rat and bovine apocytochrome b5 (Storch and Daggett).[24] From these studies it was observed that, in spite of heme removal core 2 was comparatively well maintained and the loss of secondary structure was limited to core 1 or heme binding region.

1.5 Replica Exchange Molecular Dynamics

Effective conformational sampling is one of the major challenges while conducting molecular dynamics simulations of complex proteins, which have rugged energy landscape. In such cases due to the presence of multiple minima states, proteins tend to get trapped in a local minima state, which is called kinetic trapping.[25] To solve this problem, various ensemble algorithms like multi canonical algorithm, simulated tempering and Replica exchange Molecular Dynamics (REMD) have been developed.

In Replica Exchange Molecular Dynamics several copies of the system are created at different temperature and at regular intervals exchanges of these replicas occur, which results in better sampling and also during the swapping between temperatures, bio molecules, which are confined in local minima, are released.[26] The Replica Exchange simulations can be distributed over various nodes in a cluster, thereby increasing the speed. Like any method REMD has its own drawbacks, for large bio molecular systems the number of replicas required is huge and therefore it requires several processors and utilizes large chunks of memory space.[27] Also time dependent properties cannot be calculated, as discontinuity is introduces during swapping of

replicas. Replica Exchange dynamics finds varied applications like protein folding, peptide analysis and conformation dynamics of nucleic acids. [28,29]

1.6 Apocytochrome b5 Dynamics

To understand the role of heme in cytochrome b5 protein, it is vital to study the dynamics of both heme-bound (holo form) and heme – free (apo form) of cytochrome b5. Several experimental procedures like circular dichroism (CD), fluorescence spectroscopy, NMR etc have been used to study the consequences of heme removal from cytochrome b5.[21,28-31] NMR studies by Moore and Lecomte revealed that with loss of heme, core 1 exhibited considerable secondary structure loss and increased mobility, whereas core 2 was well maintained and intact.[29,30]

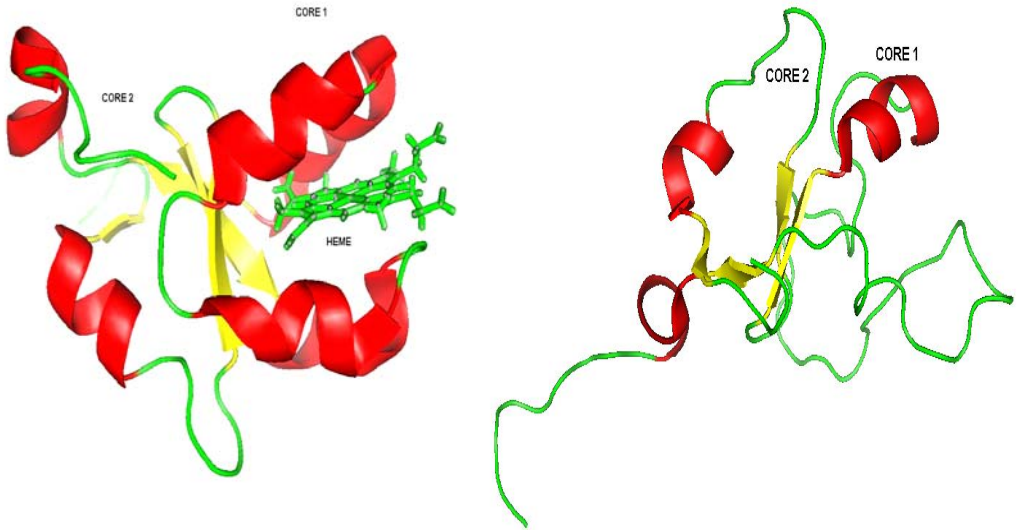


Fig.2. Rat cytochrome b5, (PDB ID: 1AW3) and Rat apocytochrome b5 (PDB ID: 1I8C)

Storch and Daggett conducted MD simulations of rat and bovine apocytochrome b5. For their simulation they used the starting structure of the holo proteins with the prosthetic heme group removed.[23] From their simulations, they found that core 1 of apo protein exhibited higher fluctuations and while the core 2 remained intact. Extended MD simulations of apocytochrome b5 were conducted using Replica Exchange method by Kuczera and co workers.[23] These simulation trajectories were about 3.1 ns in length and in order to speed up the calculations, they used the GB/SA implicit solvent model for the molecular dynamics. Benson and their group, during a recent study, concluded that upon heme removal, the heme-binding pocket of apocytochrome OM form is more compact and rigid as compared to the Microsomal form.[22]

One of the main aims of my thesis was to study the mobility and interactions of the residues, which were different in the Microsomal and OM cytochrome b5. It is also interesting to study the dynamics of the heme binding region, the hydrophobic interaction etc. Hence MD simulations were conducted to analyze the dynamics and structural differences between both the microsomal and OM forms of apocytochrome b5. The Figure 3. shows the differing residues between the Microsomal and OM proteins.

```
1I8C     4 VKYYTLEEIQKHKDSKSTWVILHHKVYDLTKFLEEHPGGEEVLREQAGGDATENFEDVGH
1B5M     4 VTYYRLEEVAKRNTAEETWMVIHGRVYDITRFLSEHPGGEEVLLEQAGADATESFEDVGH
            * ** ***  *       **   *  *** * ** ********* **** **** ******

1I8C    64 STDARELSKTYIIGELHPDD
1B5M    64 SPDAREMLKQYYIGDVHPND
            * ****  * * **  ** *
```

Fig 3. Comparison of sequences between rat Mc & rat OM form, similar residues are highlighted

1.7 Holo protein Dynamics

Several molecular dynamics simulations have been conducted on holo forms of cytochrome b5.[22,32] In order to understand the influence of specific hydrophobic residues, triple mutants and quintuple mutants of cytochrome b5 have also been subjected to MD simulations.[23]

The Replica Exchange Method was used to conduct MD simulations of rat Microsomal and rat Outer Mitochondrial proteins (shown in Fig.4.). Using the replica exchange method 12 concurrent simulations were conducted at different temperatures (320 K – 450 K) and periodically the replica conformations were swapped between the temperatures. The implicit GB/SA method was used for calculating the solvent electrostatics. The advantage of using continuum solvent model is, it isn't computer intensive, thereby allowing us to perform longer simulations of proteins.

The profiles of RMSF (Root Mean Square Fluctuation), number of native contacts and number of hydrogen bonds in secondary structures can be correlated to the stability of the proteins.Hence the replica exchange results can be compared to the results obtained from NH exchange and thermodynamic stability experiments.

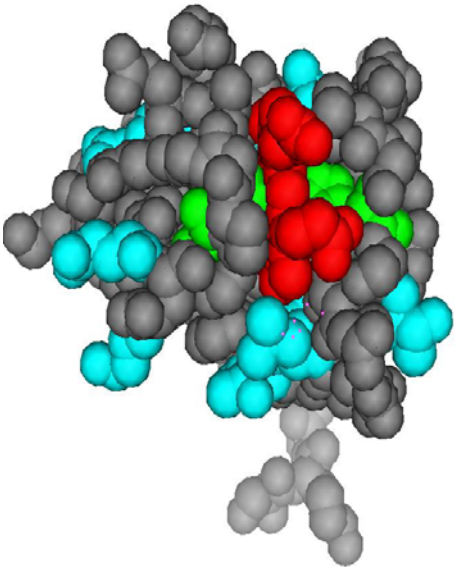


Fig .4. 1AW3 – Microsomal holo protein

Red – Heme, Green – HIS 39, HIS 63

Cyan – (L32, K34, E37,R47,G52, N57,T65, T65, L70, S71, T73)

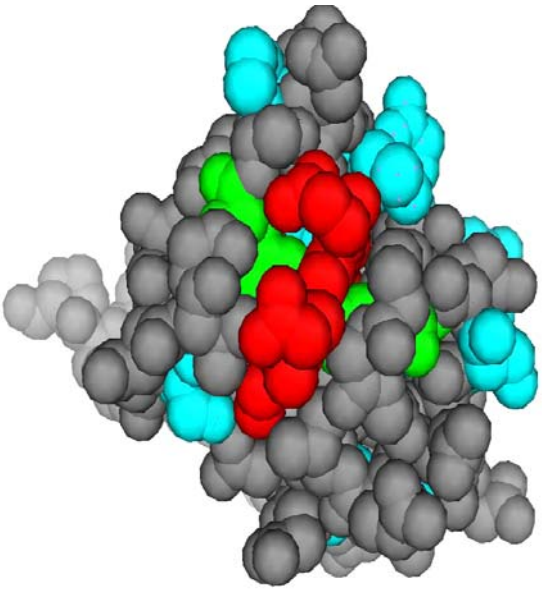


Fig.5. 1B5M – Outer Mitochondrial holo protein

Red – Heme, Green – HIS 39, HIS 63

Cyan – (I32, R34, S37, L47, A52, S57, P65, M70, L71,Q73)

2. SIMULATION PROCEDURE

2.1 MD SIMULATIONS

The Molecular Dynamics Simulations of the apocytochrome b5 protein were conducted and the simulations were 20 ns long. The starting structures were obtained from the Protein Data Bank and consisted of both X-ray , NMR structures deposited in the database.

The various files used were as follows

1. Apocytochrome b5: The microsomal apo proteins simulations were conducted using the starting coordinates of 1I8C, the average NMR structure calculated from an ensemble of 20 structures. Along with the average structure, the molecular dynamics of two more conformers were done using the collection of 20 structures 1I87.
2. Rat Microsomal cytochrome b5: The MD simulations of Microsomal cytochrome b5 was performed with the starting structure 1AW3 minimized average NMR structure of rat cytochrome b5
3. Rat outer mitochondrial cytochrome b5: The molecular dynamics of the OM form was conducted using pdb file 1B5M, the X-ray, structure with a resolution of 2.70 A^{0} .

   The simulations were conducted using CHARMM version 31 and the parameters and the calculations were conducted using the version 22-protein topology and parameters.[34,35]

The general procedure for conducting molecular dynamics simulations is described below

2.1.1 Initialization: The initial coordinates are obtained from the pdb and the trailing water molecules are deleted. After this using the HBUILD module of CHARMM, proton coordinates are constructed.[34] In case of the apocytochrome OM form since no starting NMR structure or X-ray structure was available, the starting coordinates of its Microsomal apo protein is taken and specific residues, which were different in the OM form, were replaced to construct the starting structure of apocytochrome b5 OM protein.

2.1.2 Addition of ions: In order to neutralize the protein, and to mimic the biological environment it was vital to maintain the ionic strength close to the concentration of 0.15 M. Hence additional ions were added using the program SOLVATE (version 1.1).[34] The ionic strength is expressed as total concentration of all ions present in the solution. Based on our ionic strength, we can calculate the total number of ions required and we also know the total charge of the system; hence we find out the individual number of positive ions and negative ions and add it to our system. During the rat apo cytochrome simulation 22 ions were added of which 16 Na^+ and 6 Cl^- . In case of Microsomal apo cytochrome b5, the overall charge was – 8.000 and hence total of 56 ions were added (32 Na^+ and 24 Cl^- ions).

2.1.3 Solvating the protein: The system to be simulated is put into a space filling box, which is surrounded by copies of itself. We use a truncated octahedral space filling unit cell since its much closer to a sphere, thereby requiring fewer solvent molecules to fill the cell. The Octahedral box for the Mc form, was obtained by cutting of the edges of a 65-Angstrom cube, whereas the apo cytochrome b5 required a larger water box of about 85 Angstroms. The octahedral water box was equilibrated prior to the addition of the protein. The water molecules were described using the TIP3P model and 3730 water molecules were added to the system. The apo cytochrome b5 had a larger box size and hence around 9272 water molecules were added. The water molecules, which were overlapping with the protein and ions, were removed and before the dynamics the water molecules in Mc cytb5, OM cytb5 , Apo Mc and Apo OM forms were 4348, 4348 ,4250 and 3845 respectively.

2.1.4 Equilibration of the system: The first part of the equilibration step was constrained minimization wherein the protein molecule was fixed and the rest of the system consisting of ions and water molecules were subjected to 100 steps of Adopted basis Newton Raphson energy minimization. Following the minimization they were equilibrated in order to eliminate close contacts and also eliminate geometric strains in the solvent. Later the entire system along with the protein molecule was subjected to 100 ps of equilibration.

2.1.5 Molecular Dynamics simulations:

MD simulations of both the Mc and OM form of apo-protein were conducted using CHARMM and the length of the trajectories was about 20ns. The MD simulations of the holo proteins were also conducted and the length of the trajectories was about 15 ns. The MD simulations were conducted at constant pressure of 1 atm and constant temperature of 300K. The constant temperature was maintained using the Hoover thermostat and constant pressure using the Langevin piston method. The image module of CHARMM was used to apply periodic boundary conditions and the leapfrog integrator with a time step of 2 femtosecond was used for the dynamics. Holonomic constraints were placed on the bond lengths using SHAKE algorithm. In order to adjust the inaccuracy in calculation of long-range interactions introduced by cutoff radius, the Particle Mesh Ewald (PME) method was used to calculate the electrostatic interactions. The ewald parameter κ was set to 0.34 $Å^{-1}$ and the grid spacing parameter was set to 64, the cut off distance was set to 12 Å .

The trajectories were saved and the coordinated were written to the output file every 250 steps. The mpi version of CHARMM was used to conduct these simulations and they were conducted at the bioinformatics cluster at Information & Telecommunication Technology center, Kansas.

2.2. REPLICA EXCHANGE MOLECULAR DYNAMICS

2.2.1 Background

The replica exchange molecular dynamics is a generalized ensemble algorithm, wherein non-interacting replicas of the system are created at different temperatures and the replicas are exchanged based on the Metropolis criterion. The basic idea is to increase the conformational sampling, as replicas move from one temperature bath to another; conformational changes possible at higher temperature migrate to lower temperature thereby increasing sampling efficiency.[26]

The Replica exchange method was initially developed to conduct Monte Carlo simulations and later extended to perform MD simulations. Hence the replicas are exchanged based on the Metropolis criterion given by

$\mathbf{p = 1;} \qquad \mathbf{\Delta \leq 0}$

$\mathbf{p = \exp(-\Delta);\ \Delta > 0} \qquad \mathbf{where\ \Delta = (1/\beta_i - 1/\beta_j)(Ej - E_i)}$

The two temperature baths between which the exchange occurs, is represented by i,j in the above equation and $\beta_i = 1/kT_i$, $\beta_j = 1/kT_j$. The total number of replicas required is determined based on the number of atoms in the system. The temperature distribution is selected based on the fact that in order to exchange replicas, the potential energy ranges between consecutive temperatures must overlap.[37] This overlap determined the swap acceptance probability. Usually a probability of around 0.2-0.4 is considered to be optimum for obtaining efficient sampling.

The Replica Exchange Molecular Dynamics was conducted using the MMTSB tool set, consisting of several Perl packages and scripts, which assist in performing MD simulations and analysis.[37] The starting structures were obtained from PDB files of 1AW3 for Microsomal cytochrome b5 protein and 1B5M for Outer Mitochondrial protein. The REX simulations were conducted on the bioinformatics fusion cluster and about 12 replicas were simulated for each holo protein. The distribution of temperatures was decided based on the energy profiles of the system and the distribution, which gave us an optimal overlap was selected. The 12 replicas were simulated at 320 K, 330 K, 340 K, 351 K, 362 K, 373 K, 385 K, 397 K, 410K, 422 K, 436 K and 450 K. The following picture shows us the energy profiles at different temperatures for the holo Microsomal cytochrome b5.

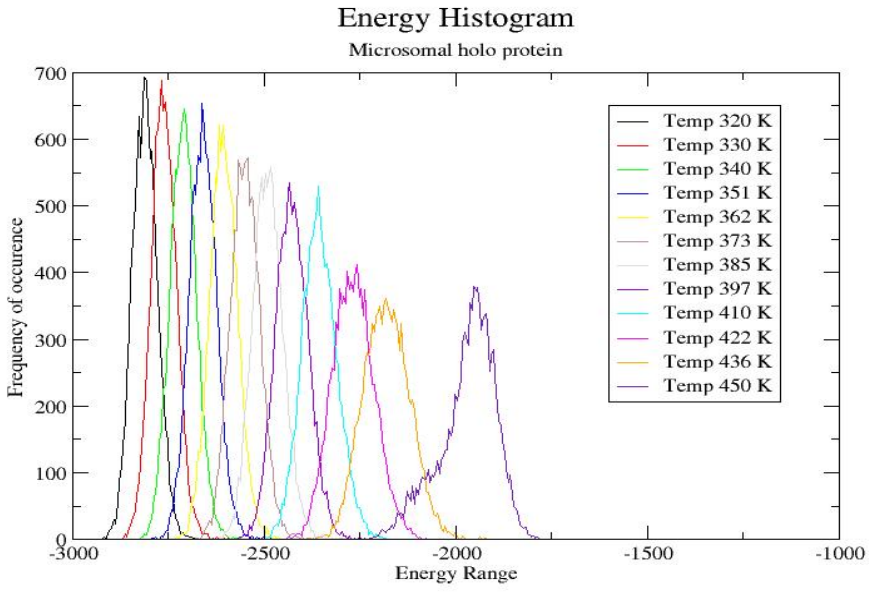


Figure 2.1 : Energy distribution for rat Mc b5 proteins.

The REMD simulations of Microsomal and OM proteins consisted of 10935 cycles. In each cycle the replicas were propagated using 200 dynamic steps with a time step of 2 femtoseconds. Overall the length of the trajectory was 10935 * 200 * 0.002 = 4.37 nano seconds. The CHARMM 22 force fields were used during the simulation and implicit GB/SA solvent model was used. The implicit solvent model represents the solvent as a continuous medium and this is very useful in solvating large bio molecular systems. The GB/SA model uses an approximation of the Poisson Boltzmann equation, which describes the electrostatic interactions. The total solvation free energy is represented as a sum of the solvent-solvent free energy (G_{cav}), solute-solvent Vander Waals energy term G_{vdW} and solute-solvent electrostatic free energy G_{pol} .

$$G_{solv} = G_{cav} + G_{vdW} + G_{pol}$$

The G_{pol} term is obtained as an approximation of the Poisson Boltzmann equation & is given by

$$G_{pol} = -166.0\left(1 - \frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{\sqrt{r_{ij}^2 + b_i b_j \exp\left(-\frac{r_{ij}^2}{4 b_i b_j}\right)}}$$

where the summation is over all pairs of atoms (i and j). The partial charges of the atoms are represented as q , the r_{ij} is the separation between the atoms i and j , ε is the dielectric constant of the medium and the Born radii of atoms is represented as b_i .[39]

The solvent solvent free energy term and Vander Waals term are represented as a linear function of the solvent accessible surface area

$$G_{cav} + G_{vdW} = \sum_{k=1}^{N} ASP(k)ASA(k)$$

where ASP(k) is the solvation parameter of the atom k and ASA(k) is total solvent accessible surface area of atom k.

3. ANALYSIS OF PROTEIN DYNAMICS

3.1 Protein Structure Analysis

The merged trajectories of the MD simulation were constructed using CHARMM and structure analysis was conducted on both forms of apocytochrome b5. The following abbreviations are used throughout the analysis chapter, microsomal apoprotein from average NMR structure – 1I8C (Apo Mc), microsomal apoproteins from ensemble of 20 structures – 1I87 (Apo Mc1 and Apo Mc2), OM protein modified from 1I8C (Apo OM1), OM proteins modified from the ensemble of 20 structures (Apo OM1 and Apo OM2).

Since we had constructed the OM apoprotein by replacing about 28 residues, the results obtained from apo OM and apo OM1 simulations weren't consistent with previous results, however the apo OM2 simulation results showed lower RMSD values and calculation of RMSF, radius of gyration and secondary structure results were all consistent to experimental results, hence mainly apo OM2 results have been discussed in the results section. The Table 3.1 shows the backbone RMS Deviations between the trajectory average and starting experimental structures. The RMS Deviations observed in holoproteins have been reported in brackets.[22] The RMSD values of microsomal apoproteins ranged between 3.2 Å (Apo_Mc1)– 6.30 Å (Apo_Mc2). The RMS Deviation of the Apo_Mc form was 3.70 Å. The RMSD values of the apoproteins were almost thrice those of the Microsomal holoproteins.

Table 3.1 : Backbone RMSD of trajectory average structure from experimental structure.

Proteins	RMSD of backbone atoms (Å)
Apo_Mc	3.70 (1.1)
Apo_Mc1	3.20
Apo_Mc2	6.30

Table 3.2: Backbone RMSD around trajectory average

Proteins	RMSD of backbone atoms (Å)
Apo_OM	7.00 (1.2)
Apo_OM1	6.80
Apo_OM2	4.00

The RMS deviation of the OM Apocytochrome b5 was in the range of 4 Å – 7 Å . Among the apo OM simulations the Apo OM 2 simulation results were comparable to the Microsomal simulations.

The Figures 3.1 and 3.2 shows the RMSD values of residues 4-85 for both the Mc and OM proteins. In the apo microsomal protein increased RMS Deviations were observed in the regions 17-20, 45-50 and 60-65, and the Apo_Mc form exhibited the lowest RMSD, whereas in the apo OM protein, the apo OM 2 had the lowest RMSD.

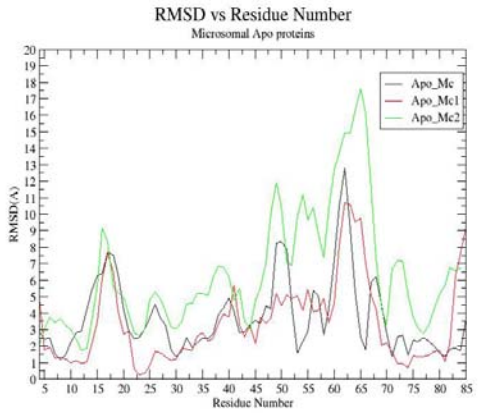


Figure 3.1: RMSD of Microsomal proteins

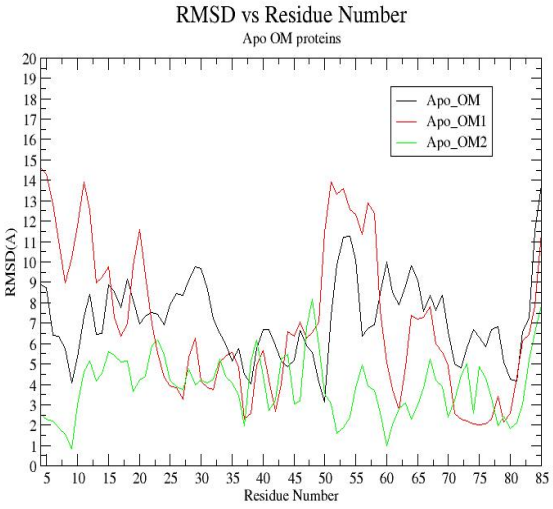


Figure 3.2 : RMSD of OM proteins

3.2 Structure Analysis of Core 1 & Core2

Core 1 defines the heme binding pocket and comprises of the residues 32:76, while Core 2 comprises of the residues 4:32 and 76:85.The C_α RMS Deviation from the starting structure was observed as a function of time for all the apo proteins. The RMSD of core 1 in Microsomal proteins was between 4.8-8.7 Å, whereas in the OM proteins it ranged between 4.2-6.5 Å. The RMS Deviations of core 2 in apo Mc proteins was in range of 1.6-3.2 Å and in apo OM proteins it was in the range of 4.5-8 Å. In case of all the microsomal apo proteins the average RMSD value for core 1 were larger than core 2 & Figure 3.1 shows the difference between the cores. However in some of the OM proteins the core 2 has a higher RMS deviation as compared to its core 1 and this can be explained due to the excessive mobility of the N termini and C termini.

Table 3.2 : Average RMSD of core 1 and core 2

Proteins	Avg RMSD - Core 1 (Å)	Avg RMSD – Core2 (Å)
Apo_Mc	4.80	1.60
Apo_Mc1	4.60	3.20
Apo_Mc2	8.70	2.30

Apo_OM	6.50	4.90
Apo_OM1	4.80	8.0
Apo_OM2	4.20	4.50

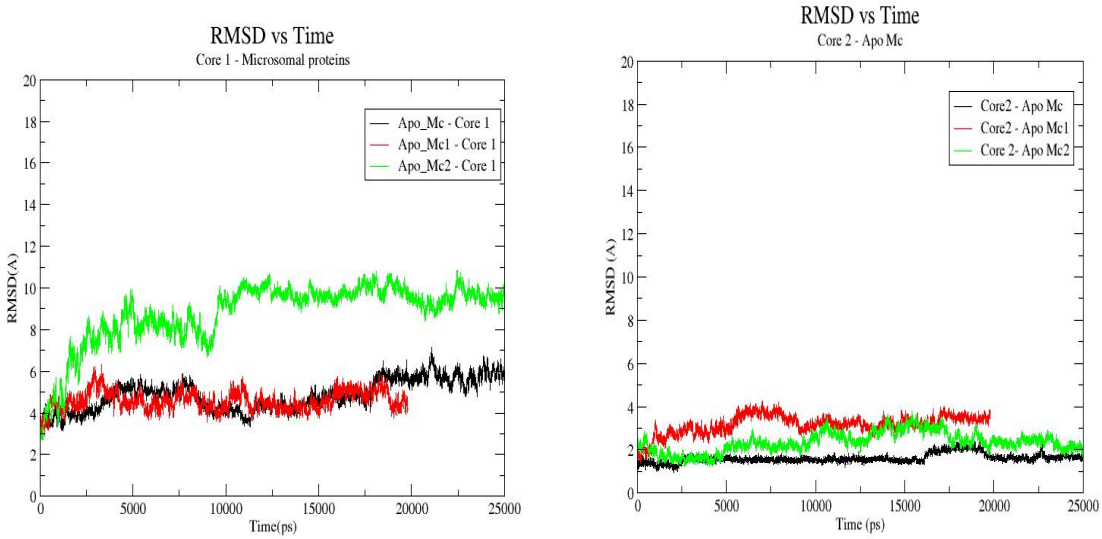


Fig 3.3 RMSD vs. time for Core 1 and Core 2 – Microsomal apo proteins

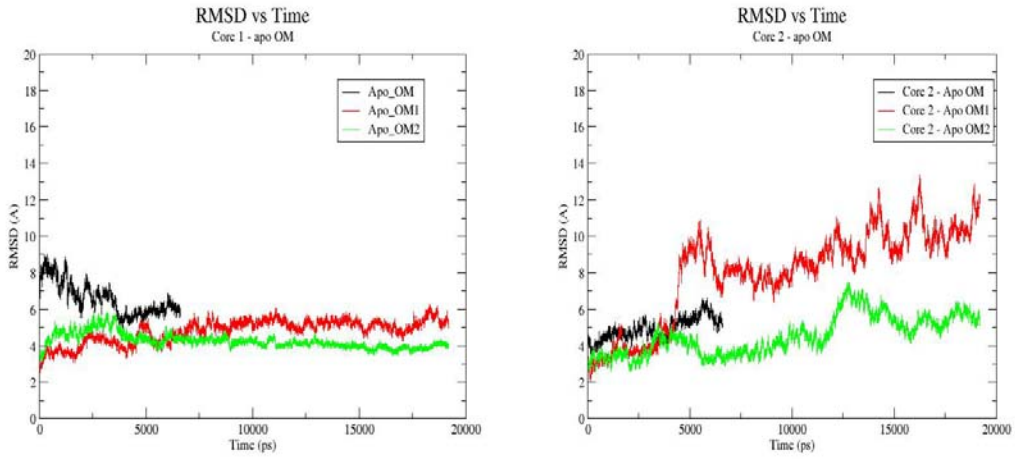


Fig 3.4 RMSD vs. Time for Core 1 and Core 2 – Apo OM proteins

3.3 RMS Fluctuations

The RMS Fluctuations of both Microsomal and OM apoproteins were calculated from the trajectory average. In case of the Microsomal apoproteins, the lowest fluctuation was in the Apo_Mc form (1.8 Å) and highest fluctuation was in the Apo_Mc2 (3.5 Å). In the OM protein structures the Apo_OM1 form had the highest fluctuation (4.1 Å) and the Apo OM form had the lowest fluctuation (3.0 Å). Similar to the RMSD values observed the RMS fluctuations of the apoproteins were almost twice that of the holoproteins, the holoprotein RMSF values are given in the bracket.

Table 3.3 : RMSF of backbone atoms

Proteins	RMSF of backbone atoms From Avg Trajectory Structure (Å)
Apo_Mc	1.80 (0.7)
Apo_Mc1	2.00
Apo_Mc2	3.50
Apo_OM	3.00 (0.9)
Apo_OM1	4.10
Apo_OM2	3.10

Apart from measuring the mobility of the entire protein we also compared the RMS Fluctuations of core 1 and core 2, as the RMSD values aren't helpful to analyze about the mobility. The results obtained are consistent with the previous Molecular dynamics results. In case of the microsomal proteins the core 1 was more mobile than the core 2 and this is shown in the following Figure 3.5.

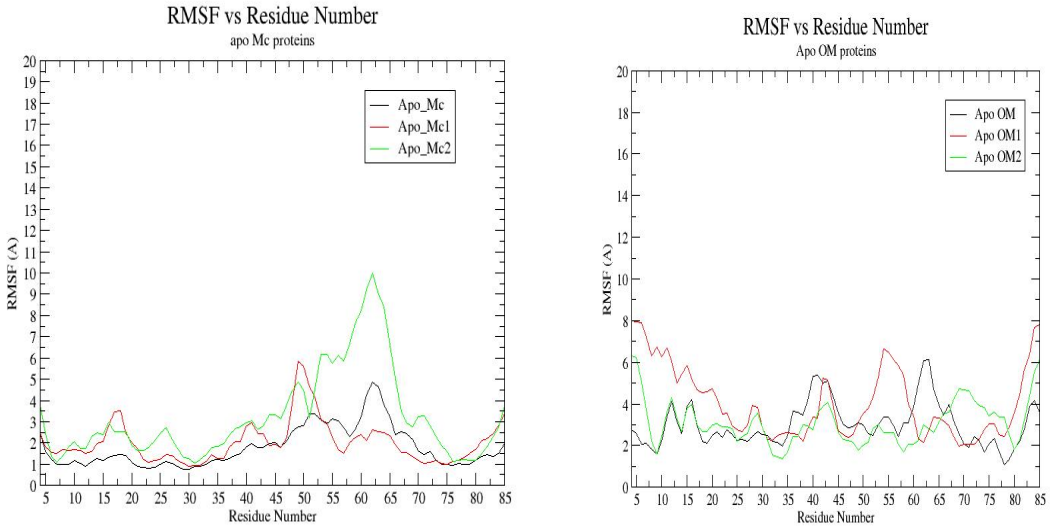


Figure 3.5: Backbone RMSF of Core 1 and Core 2 - Microsomal Proteins

A comparison of the RMS Fluctuations between the Microsomal and OM apoproteins show that, overall the Mc apoprotein was more flexible than the OM apoprotein.

Also the heme binding pocket in the apo OM protein shows lower conformational mobility when compared to the Microsomal protein as reflected by low RMSF values. This is consistent with the idea that even in the absence of the heme group the hydrophobic interactions in the two conserved hydrophobic patches (Ile 25, Phe 58 and Leu 71) residues, (Ala 18, Ile 32, Leu 36 and Leu 47) stabilize the apo OM core 1.[21,40]

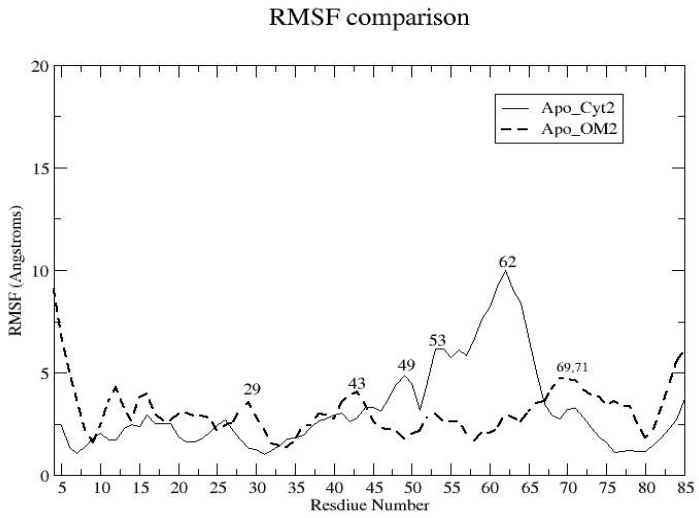


Fig 3.6: Comparison of RMSF between Mc and OM apoproteins

The RMS fluctuations of the histidine ligands (His 39, His 63) are tabulated in the table 3.4 and the histidine ligands are located in loops that connect consecutive helices and RMSF of these regions are analyzed. RMS Deviations of these regions are compared. The His 63 residue fluctuations are larger than the His 39 residue fluctuations in all of the microsomal apoproteins. However this kind of difference in histidine fluctuations wasn't observed in the Apo OM2 protein.

Table 3.4: RMSF of Histidine 63 and 39

Proteins	RMSF (Å) His 39	RMSF (Å) His 63	RMSD (Å) His 39	RMSD (Å) His 63
Apo_Mc	1.50	4.60	4.30	9.10
Apo_Mc1	2.10	2.50	3.90	10.50
Apo_Mc2	2.70	8.90	6.70	14.90
Apo_OM	4.20	6.10	5.70	8.80
Apo_OM1	2.80	2.80	5.00	4.80
Apo_OM2	2.90	2.80	6.20	3.10

The RMS Deviations of the histidine loops are listed in the Table 3.5. The deviation of the Histidine 63 loop is higher in case of all the Microsomal proteins and from the RMSF data we can conclude that this loop is very mobile as compared to the His 39 loop. These results are in agreement with NMR studies, which suggested that the His 63 residue was more mobile and heme association studies which, predicted that the His 63 loop had higher flexibility as compared to the His 39 loop.[28]

Table 3.5: RMSD of histidine loops

Proteins	RMSD (Å) - Loop 38 - 43	RMSD (Å) – Loop 60 -65
Apo_Mc	4.50	9.40
Apo_Mc1	4.70	10.20
Apo_Mc2	6.50	15.80
Apo_OM	6.50	9.40
Apo_OM1	5.53	6.30
Apo_OM2	7.40	4.00

Another interesting result obtained from the RMSD of individual residues is the β5 strand region covering the residues (51 – 55) had higher RMS Deviations in case of the Microsomal proteins as compared to the OM proteins. The RMS Deviations of this loop are given in Table 3.6.

Table 3.6: RMSD of loop (51-55)

Proteins	RMSD (Å) Loop (51 – 55)
Apo_Mc	4.60
Apo_Mc1	5.70
Apo_Mc2	9.80
Apo_OM	10.60
Apo_OM1	13.09
Apo_OM2	3.50

3.4 Radius of Gyration:

The radius of gyration is the mass weighted distance of each atom from the molecular center-of-mass and this would give us an estimate of the size of protein. In our case however in addition to the entire protein size we are interested in the core sizes of the Microsomal and Outer Mitochondrial proteins. The equation 3.1 shows the formula with which radius of gyration is calculated.

$$R_g = \left(\frac{\sum_i \|r_i\|^2 m_i}{\sum_i m_i} \right)^{\frac{1}{2}} \quad \text{.....Eq 3.1}$$

Where m_i is the mass of atom i and r_i is the position of the atom with respect to the center of mass of molecule. Table 3.7 compares the radius of gyration of different

apocytochrome b_5 proteins. The Apo_Mc1 has the lowest core size of about 11.36 Å and in the Microsomal proteins the largest radius of gyration was found in the Apo_Mc2 structure. In the holo proteins, from previous simulation done by Cheng et al, we find the core 1 size to be around 10.5 Å.[22] The apocytochrome proteins show larger radius of gyration values as compared to holo proteins. In the OM form, the largest radius of gyration was observed in the Apo OM protein (12.88 Å) and the Apo OM2 protein had the lowest radius of gyration value of 11.51 Å. As observed from the table the radius of gyration of the apo proteins is about the same for both the isoforms and ranges between (11.36 – 12.88 Å).

Table 3.7: Radius of Gyration of heme binding core 1

Proteins	Radius of Gyration Core 1 (32-76) - Å
Apo_Mc	11.53
Apo_Mc1	11.36
Apo_Mc2	12.88
Apo_OM	12.61
Apo_OM1	11.71
Apo_OM2	11.51

3.5 Secondary Structure

In order to analyze the preservation of secondary structure, the RMSD of the helices and β sheets were calculated. To supplement the RMSD results, the hydrogen bond distance between backbone atoms N (i + 4) and O (i) were calculated for three helices. A hydrogen bond was defined as having a distance of 3.6 Å and the number of hydrogen bonds present in the trajectory average was calculated. From this the fraction of bonds, which were within 3.6 Å, was calculated and represented as a percentage.

In general the secondary structure of protein was well maintained for the microsomal apoprotein as compared to Outer Mitochondrial apoprotein. Helices α1 (8-13) and α2 (32-36) retained helicity and were very stable for the Apo Mc proteins as reflected by the RMSD values which remained below 0.25 Angstroms. Unlike the cytochrome b5 holo protein, in almost all the apoproteins there were only 3 helices and 4 β sheets. As discussed before the helix 3 (which corresponds to α5 in holo proteins) is comparatively disrupted and similar results were observed during the MD simulations of Storch and Daggett.[23] During their MD simulations they observed that the helix α 3 (69-73), became quite disrupted at the N terminus. In case of the β sheets, for most of the microsomal apoproteins the RMSD values were below 0.4 Å, however the deviations were higher in the OM apoproteins and had a wide range of 0.3 – 1.0 Å. While looking at the individual β sheets we observe that β2 and β4 strands were well

maintained and were quite stable as shown by the RMSD values. The β1 strand (residue 5-7) and β4 strand (77-79) fluctuated a lot in the OM apo forms and this might be due to the excessive movement of the N and C termini thereby resulting in closer contacts with protein core.

Table 3.9: RMSD of helices for all apocytochrome proteins

Proteins	Helix 1 8 –13 (Å)	Helix 2 32 – 36 (Å)	Helix 3 69 - 72 (Å)
Apo_Mc	0.11	0.24	1.36
Apo_Mc1	0.11	0.18	0.65
Apo_Mc2	0.13	0.23	2.10
Apo_OM	0.59	0.69	0.60
Apo_OM1	1.37	0.66	1.29
Apo_OM2	0.77	0.78	0.97

Table 3.10: RMS Deviations of β strands for all Apocytochrome proteins

Proteins	β 1 5 – 7 (Å)	β 2 22 – 2 4 (Å)	β 3 29 – 31 (Å)	β4 77 – 79 (Å)
Apo_Mc	0.34	0.28	0.42	0.14
Apo_Mc1	0.76	0.25	0.18	0.22
Apo_Mc2	0.40	0.30	0.28	0.19
Apo_OM	0.55	0.30	0.45	0.78
Apo_OM1	0.98	0.38	0.42	0.76
Apo_OM2	0.54	0.37	0.20	0.69

The hydrogen bond N…O distances were below 3.6 Å for both helix 1 and helix 2 in all the microsomal proteins. Since the hydrogen bond is represented as deviation percentage, for both these helices consisting of residues 8-13 and 32-36 respectively, the percentages were above 90% for most of the cases. In contrast the helix 3 (residues 69 – 45) was distorted and had increased N……O distances of above 5.00 Å. The Apo Microsomal protein (from 1I8C) was an exception wherein initial parts of the helix had about 90 % H-bonds, however even this helix was interrupted in between and fluctuated a lot. In case of the OM apo proteins the helices 1 was well maintained with the exception of hydrogen bonds (observed at the starting of the

helices). The helix 2 was a bit distorted in all the OM proteins and this might be due to the absence of stabilizing heme peptide interaction, which is present in the holo protein. In the Apo OM 2 protein, the helix 3 (69 – 72) wasn't maintained and this region exhibited large RMS Deviations.

Table 3.11: Percentage of H-bonds for Helix 1 (8-13)

Proteins	12 N 8 O	13 N 9 O	14 N 10 O	15 N 11 O
Apo_Mc	95.30 %	97.90 %	98.70 %	97.30 %
Apo_Mc_1	86.60%	97.60 %	92.50 %	44.30%
Apo_Mc_2	76.70 %	96.30 %	92.80 %	51.30 %
Apo_OM	-	-	81.10 %	94.80 %
Apo_OM1	-	-	79.90 %	85.07 %
Apo_OM2	-	-	96.10 %	92.90 %

Table 3.12: Percentage of H-bonds for Helix 2 (32-36)

Proteins	36 N 32 O	37 N 33 O	38 N 34 O	39 N 35 O
Apo_Mc	98.50 %	94.80 %	80.80 %	94.2 %
Apo_Mc_1	99.70 %	68.90 %	94.10 %	99.8 %
Apo_Mc_2	99.50 %	51.50 %	72.20 %	99.8 %
Apo_OM	-	29 %	81.70 %	52.5 %
Apo_OM1	41 %	51 %	-	50.1 %
Apo_OM2	-	-	1.0 %	-

Table 3.13: Percentage of H-bonds for Helix 3 (69-73)

Proteins	73 N 69 O	74 N 70 O	75 N 71 O	76 N 72 O
Apo_Mc	90.50 %	84.2 %	21 %	-
Apo_Mc_1	25.92 %	18.38 %	-	43.18 %
Apo_Mc_2	-	-	12 %	-
Apo_OM	-	-	94.21 %	94.37 %
Apo_OM1	64.20 %	59.47 %	87.80 %	98.74 %

3.6 Solvent Accessible Surface Area

The Solvent accessible surface areas for all of the apoproteins have been tabulated and they are compared with the SASA values obtained for the holoproteins. Along with the overall solvent accessible surface area, the solvent exposure for core 1, core 2, specific hydrophobic residues and differing residues in core 1 etc have also been tabulated. The core 1 of the apoprotein is more exposed to the solvent as compared to the core 2, due to the missing heme group. Though there isn't much difference among the apo proteins SASA values, it is observed that the SASA values of OM apoproteins in core 1 are higher than the microsomal apoproteins.

Table 3.14: Trajectory Average SASA values

Proteins	Core 1	Core 2
Apo_Mc	3656 ± 108	2197 ± 142
Apo_Mc1	3483 ± 144	2511 ± 145
Apo_Mc2	3827 ± 159	2401 ± 108
Apo_OM	3780 ± 401	3385 ± 139
Apo_OM1	3580 ± 142	3959 ± 343
Apo_OM2	3687 ± 352	3330 ± 180

The solvent accessibility of the lone tryptophan (residue 22) present in core 2, is calculated and reflects the structural integrity of core 2. The SASA values of Trp 22 in Microsomal proteins are comparable to the values of holo proteins (Holo proteins – 9 % solvent exposure). The Solvent accessibility surface areas of the axial histidine residues are tabulated below. When the solvent accessibilities of histidines were compared it was found that the His 63 residue was more exposed to the solvent as compared to His 39 residue in both isoforms of apoproteins.

Table 3.15 : Percentage of solvent exposure in Apo proteins

Proteins	Trp –22	His -39	His - 63
Apo_Mc	9.5 %	38 %	80 %
Apo_Mc1	13 %	27 %	67 %
Apo_Mc2	23 %	36 %	71 %
Apo_OM	15 %	70 %	88 %
Apo_OM1	37 %	65 %	40 %
Apo_OM2	34 %	40 %	84 %

The Solvent accessibilities of the two hydrophobic networks (Res 18, 32, 26 and 47) and (Res 25, 58, 71) were calculated and tabulated in Table 3.16. As observed in holo proteins, the least SASA value is observed in case of Residue 32. Also similar to results from holo protein dynamics the Residue 18 is more exposed to the solvent in the Microsomal proteins, as the hydrophobic network is broken due to presence of Serine 18 in the Microsomal protein. When compared to holoproteins the residue 47 is more exposed to solvent and this might be due to the absence of the heme group. There is no significant difference in the solvent exposure of hydrophobic patch 1 between the Microsomal and OM apo proteins.

Table 3.16: SASA values of hydrophobic patch 1 (Residues 18, 32, 36 and 47)

Proteins	Residue18	Residue 32	Residue 36	Residue 47
Apo_Mc	67 %	0.5 %	20 %	49 %
Apo_Mc1	90 %	2 %	11 %	26 %
Apo_Mc2	90 %	4 %	36 %	22 %
Apo_OM	57 %	6 %	38 %	68 %
Apo_OM1	61 %	4.5 %	43 %	37 %
Apo_OM2	60 %	6 %	52 %	84 %

Since all of the residues belonging to the second hydrophobic patch (residue 25, 58 and 71) are in contact with the heme group, in the absence of heme the solvent exposure increases as opposed to the cytochrome b5 holoproteins. Residue 58 had a higher solvent exposure in the Microsomal protein (48 %) than in the Apo OM2 (5 %).

Table 3.17: Solvent exposure percentage

Proteins	Residue 25	Residue 58	Residue 71	Residue 35
Apo_Mc	8 %	16 %	12 %	10 %
Apo_Mc1	3 %	7 %	3 %	12 %
Apo_Mc2	8 %	48 %	39 %	9 %
Apo_OM	43 %	84 %	22 %	27 %
Apo_OM1	31 %	84 %	43 %	28 %
Apo_OM2	3 %	5 %	43 %	35 %

3.7 Inter residue distances

The average distance between the Histidine Nε2 atoms of residues 39 and 63 was calculated for all the proteins. The Apo_Mc structure had the lowest value of 19.1 Å and the Apo_Mc2 had the highest distance of 26.6 Å. The Apo OM2 protein had an inter residue distance of 20.2 Å. The inter residue distance, radius of gyration, core 1 fluctuations etc indicates that the apo OM protein has a more compact heme binding pocket than the Microsomal protein. The distances between residues 18-50 were also calculated and here again the Apo OM 2 had a shorter distance of 20.7 Å as compared to the Apo Mc 2 protein (28.00 Å).

Table 3.18: Average Distance

	His 39 – His 63 Distance (Å)	Distance (Å) Residue 18-50
Apo_Mc	19.1 ± 3.0	29.40 ± 1.3
Apo_Mc1	23.3 ± 4.0	19.60 ± 4.0
Apo_Mc2	26.6 ± 5.0	28.00 ± 1.5
Apo_OM	34.17 ± 5.0	25.90 ± 3.0
Apo_OM1	21.1 ± 3.0	27.6 ± 5.0
Apo_OM2	20.2 ± 2.8	20.7 ± 2.0

3.8 Trajectory of 20 NMR structures

Using the 20 NMR structures of apo cytochrome b_5 deposited by Falzone et al (pdb file 1I87), a trajectory was constructed and conformational analysis was performed. The RMS Fluctuations and RMS Deviations of this trajectory was compared with the 20 ns trajectory of average apo Mc protein. The 20 ns trajectory was constructed using the starting structure 1I8C, which is the average NMR structure of Microsomal apo protein. The atomic fluctuations and deviations from starting structure were compared and shown in following figures. The RMS Fluctuations of the backbone atoms were calculated and we find that the atomic fluctuations observed in the trajectory constructed from 20 NMR structures (average RMSF 1.6 Å) was almost similar to the Microsomal apcytochrome b_5 trajectory (average RMSF 1.8 Å). Specifically comparison of the Root Mean Square Fluctuations in the core 1 region (residues 32-76) showed that both these graphs were very similar and varied only in the intensity.

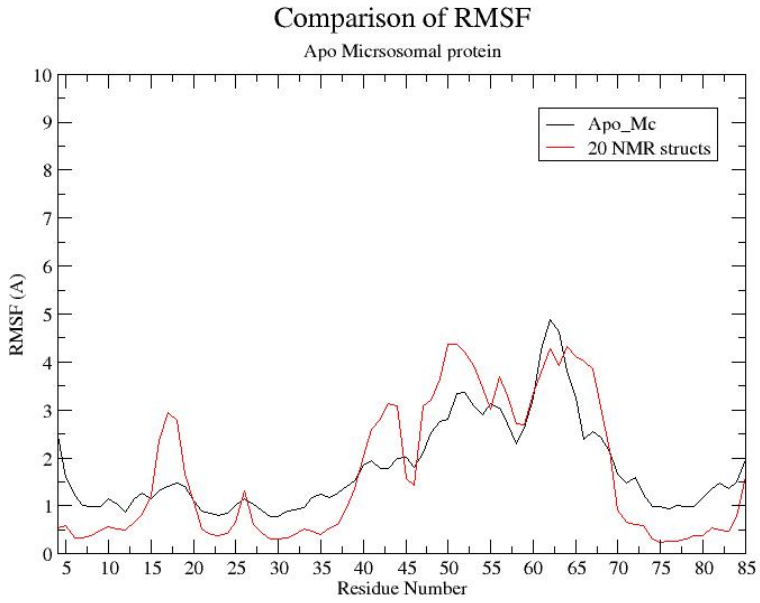


Fig.3.7: RMSF of Apo Mc and trajectory from 20 NMR structures.

A comparison of the Root Mean Square Deviations showed that the deviations were higher in case of the 20 ns average apo protein. The RMSD of 20 NMR structures trajectory from the average NMR structure was very less as expected and was about 0.84 Å, whereas the RMSD of the 20 ns trajectory was 3.4 Å. The increased RMS Deviation might have been due to fluctuations of side chain atoms in the apo protein.

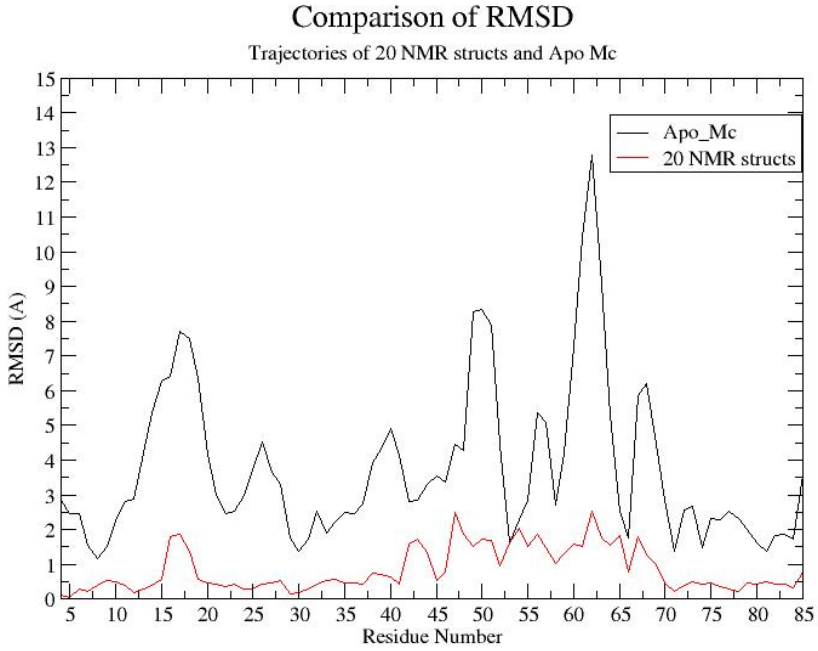


Figure 3.8: RMSD of apo Mc and trajectory from 20 NMR structures.

3.9 DISCUSSION

The structural fluctuations and dynamics of the cytochrome b5 proteins were analyzed in the absence of heme using MD simulations. The Microsomal apocytochrome b_5 simulations were conducted using the average NMR structure – PDB Id 1I8C. The OM apo cytochrome b_5 simulations were conducted using the coordinates of the Microsomal cytochrome b5 (1I8C) and by replacing all the residues, which differed in sequence.

Previous simulations of the apo forms have been conducted by removing the heme group from the holo protein structure.[24] The main inferences of heme removal from cytochrome b_5 observed by Storch and Dagett were loss secondary structure and increased RMSF values in the core1, whereas core 2 was well maintained and retained native structure.[24] The heme removal resulted in loss of secondary structure in helices α3, α4, α5 and the beta strand β5 exhibited conformational multiplicity.[24]

The simulation trajectories were about 20 ns in length and were conducted in explicit water. The apo cytochrome b_5 proteins exhibited larger deviations from the starting structure and the apocytochrome b_5 was conformationally more mobile than the holo proteins due to the absence of heme. The RMS Fluctuations of the Apo Microsomal proteins were almost twice the RMSF values of holo proteins. This is consistent with experimental and previous MD simulation results.[23,24]

The core1 comprising of residues 32:76 exhibited higher mobility than the core 2 region and this difference was more pronounced in the Microsomal apocytochrome b5 simulations rather than OM apocytochrome b5 simulations.In the OM protein higher fluctuations were observed in the C and N terminal and experimental results confirm the mobility of the α 6 (81-86) region.[41]

A comparison of the RMS Fluctuation graphs between the apo and holo Microsomal protein simulations show that the dynamics and structural fluctuations observed due to heme removal was mainly restricted to the heme binding core 1 and the core 2 was well maintained, intact. (Fig 4.16) The individual residue RMS Fluctuations was calculated and in all the apo Microsomal proteins the Histidine 63 was more mobile than the Histidine 39 residue. A comparison of the histidine loop regions (38-43) and (60-65), gave similar results and the loop region around Histidine 63 exhibited higher conformational mobility. However a similar result wasn't observed in the apo OM protein simulations and this Histidine loop region (60-65) was quite rigid. A recent study about the conformational plasticity of holo proteins using hydrogen exchange, also revealed that the region between α 4(55-60) and α5 (65-71), had lower conformational flexibility in the rat OM proteins and this was attributed due to extensive hydrophobic packing.[43]

The MD simulation trajectories were also compared to the trajectory constructed from the 20 NMR structures (PDB ID: 1I87) and it was observed that the RMSF obtained from 20 NMR trajectory closely matched with the 20 ns simulation of apo Mc protein. This is encouraging and suggests that the MD simulations were successful in observing the dynamics of the protein. However using the 20 conformers from NMR resulted in over prediction of some RMSF values as compared to MD simulation results. A comparison of the apo Mc and apo OM protein reveals that the apo OM protein shows lower conformational mobility than Mc apo protein in β5 (51-55) region and the His loop region (60-65). Conversely the Mc apo protein was more rigid than apo OM in some of the core 2 regions β4 (21-25) and β3 (27-32) region. Comparing the plots we observe that in Microsomal apocytochrome b5 simulations residues 49,52 and 62 exhibit high RMSF values, however no such peaks are observed in the OM apocytochrome b5 simulation. Overall the Microsomal apocytochrome b5 exhibited increased conformational mobility as compared to OM apocytochrome b5.

The core sizes of both the Microsomal and Outer Mitochondrial apo proteins were about the same, however as expected they were slightly greater the holo protein core sizes. Secondary structure analysis of the apo proteins reflected that the α1 (8-13) and α2 (32-36) helices were well maintained in all of the Microsomal proteins. The percentage of hydrogen bonds calculated throughout the trajectory showed that the α

5 (69-72) helix was partially distorted and similar results were obtained from the RMSD values of this region. This helix was completely disordered in the Apo OM simulation. Consistent with the observations of Storch and Daggett, the bond distances in the α6 helix were between 5 Å – 6 Å.[24] From Microsomal protein simulations, we also observe that the β 5 sheet (51-55) formally part of core 1 , was highly mobile and exhibited high RMS Fluctuations. In the OM protein (Apo OM2) the RMSF value was lower than those of Microsomal proteins and this might be due to the stabilizing side chain interaction of Ala 52 residue with the methylene group of the Gln 49 residue.[43] Due to the absence of heme group, the solvent accessibility of all the residues, which were in contact with the heme increased. In the Microsomal proteins the solvent accessibility of both the Phe 35 and Phe 58 residue increased. Another observation was consistently the solvent exposure of His 63 residue was higher than the His 39 residue in the Microsomal proteins.

The solvent accessibility of Trp 22 residue was calculated as this can compared to experimental fluorescence studies results. The solvent accessibility of Trp 22 was comparable to holo protein simulations, indicating a well-maintained core 2. The SASA values were also calculated for hydrophobic networks found in the cytochrome b5. As observed in holo proteins, the residue 32 was protected from the solvent. The solvent exposure of other hydrophobic residues was slightly increased but no significant changes were observed. The His 39 – His 63 distance was calculated in all the apo proteins and the inter residue distance was comparatively lower in the apo

OM protein and this was supplemented by core 1 radius of gyration results which showed a slightly more compact heme binding core in the Apo OM2 protein. This is consistent with experimental results, which showed that the heme-binding pocket in apo OM protein was more compact and rigid as compared to the Microsomal protein.[22]

3.10 CONCLUSIONS AND FUTURE WORK

Our Molecular dynamics simulation results are in agreement with previous inferences like that the loss of heme results in structural distortion and increased flexibility mainly in the core 1 region and the apo OM proteins have a more rigid and compact heme-binding core as compared to the Microsomal apocytochrome b5. The observation that rat Microsomal protein had a more mobile core 1 as compared to rat OM protein is in support of the experimental results that even in the absence of heme, strong hydrophobic interactions were present in the heme-binding core of rat OM. Increased fluctuations were observed in the helix ∀6 (81-86) region in case of the OM apocytochrome b5 simulation. Comparison of simulations of Microsomal apocytochrome b5 and OM apocytochrome b5 showed that overall the Mc apo protein was more flexible than OM apo protein.

3.11 FUTURE WORK

During our simulations, we found that the substitution of several residues to construct the apo OM proteins did not provide consistent results. Hence we are working on constructing a better model of the apo OM protein using the core 2 region of holo OM protein and replacing the residues in the core 1 region alone. Apart from this it is interesting to conduct MD simulations of mutated Microsomal proteins, in the beta sheet region (51-55) and in histidines loop region (60-65).

4. REMD OF HOLO PROTEINS

Using the Replica exchange method we have performed 4.37 ns simulations, which depict the temperature variation of several properties of cytochrome b_5 holo protein. Conformational analysis of the simulations has been performed and the atomic fluctuations, RMS Deviations, Solvent accessibility, percentage of native contacts and secondary structure have been tabulated. As mentioned in the methods section, the starting structure for conducting the dynamics was 1AW3 (for the microsomal protein) and 1B5M (for the OM protein). From the replica exchange results, CHARMM trajectories were constructed for 12 temperatures ranging from 320 – 450 K. Higher temperatures (i.e 400 K) have been considered in order to increase the conformational sampling during the exchange process.

4.1 RMSF of cytochrome b5

From the CHARMM trajectories, the trajectory average structures of both rat Mc and rat OM b_5 were calculated for different temperatures. The backbone RMS fluctuations were calculated for the distribution of structures at various temperatures. The corresponding melting curve is shown in Figure 4.1 and the values are tabulated in the Table 4.1. In the temperature region 320 K – 362 K, the RMS Fluctuations of the Microsomal cytochrome b5 ranged from 1.32 Å – 2.09 Å. For the OM holo protein in this temperature region, the RMSF values were between 2.03 Å – 2.23 Å. After this

in the Microsomal protein a gradual increase was observed in the RMS Fluctuations, from 2.09 Å - 6.68 Å, which corresponded to the temperature range 373 K – 436 K.

The RMSF values of the rat OM protein fluctuated around 2.00 Å and the backbone fluctuations were stable till 436 K, where it increased to 4.47 Angstroms. In both the Mc and OM protein, when the temperature increased to 450 there was a steep increase in the RMS Fluctuation to about 14.33 Å in Mc form and 16.26 Å in the OM.

4.1 Table: RMSF of backbone atoms around trajectory average.

Temperature (K)	Rat_Mc (Holo protein) Å	Rat_OM (Holo protein) Å
320	1.32 ± 0.42	2.03 ± 0.76
330	1.46 ± 0.47	1.99 ± 0.75
340	1.65 ± 0.55	2.04 ± 0.78
351	1.86 ± 0.64	2.21 ± 0.85
362	2.09 ± 0.72	2.23 ± 0.9
373	2.29 ± 0.78	2.13 ± 0.98
385	2.82 ± 0.96	2.00 ± 1.1
397	3.85 ± 1.5	1.97 ± 1.2
410	5.57 ± 2.4	2.10 ± 1.3
422	6.24 ± 2.5	2.95 ± 1.6
436	6.68 ± 2.6	4.47 ± 2.4
450	14.33 ± 2.7	16.26 ± 4.2

From the plots of RMSF vs Temperature constructed, we observe that as the temperature increases the RMSF values of the rat OM holoprotein is quite stable until 422 K, whereas in the rat Microsomal protein there is a steady increase in the temperature. This agrees with the experimental thermal denaturation results, which show that the Outer Mitochondrial protein is more stable than the Microsomal protein.[22] Also the Rat OM protein is more flexible than the Microsomal protein at very high temperatures, as observed from the RMSF value at 450 K. Comparison of melting curves show that the Microsomal rat protein have a two stage unfolding, as reflected by the bump in the plot.However the rat OM protein doesn't show any bump and seems to unfold in a single step.

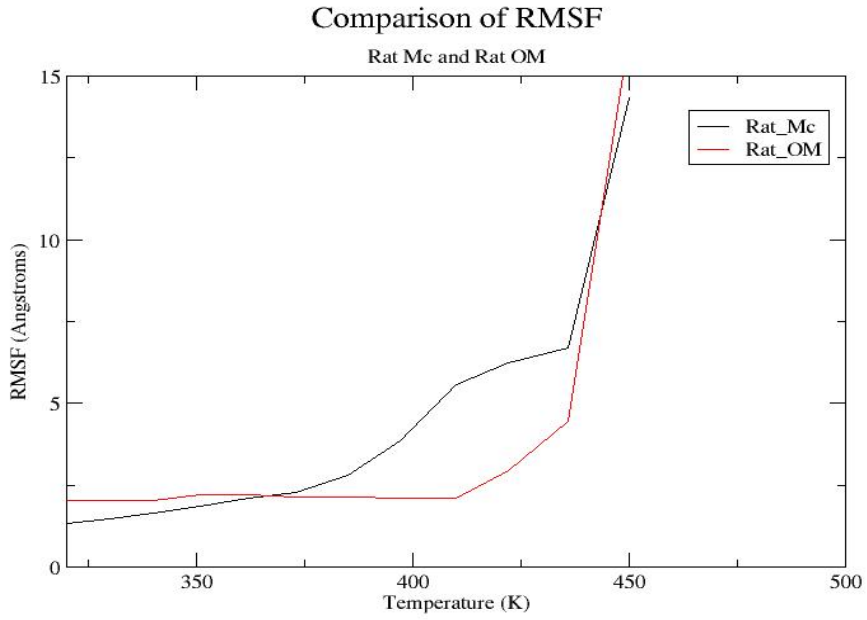


Figure 4.1: Temperature vs RMSF of rat Mc and rat OM proteins

4.2 RMSD of backbone atoms

The RMS Deviations of the Microsomal protein were in the range of 0.90 Å – 1.05 Å between the temperatures 320 K – 385 K. As observed in the RMSF values the RMS Deviations steadily increased from 1.40 Å at 397 K to 3.94 Å at 436 K, after which there was a sudden increase to 16.07 Å at 450 K. The RMS Deviations of rat OM protein ranged between 0.83 Å – 0.86 Å in the 320 – 385 K temperature region. Beyond this temperature the RMS Deviation increased from 0.94 Å at 397 K to 2.47 Å at 436 K. Then there was a steep increase to 12.84 Å at the temperature 450 K. This steep increase in temperature in RMS Deviations and RMS Fluctuations suggest that both the holo proteins undergo thermal unfolding at higher temperatures. In order to confirm this we also looked at the secondary structures of the proteins.

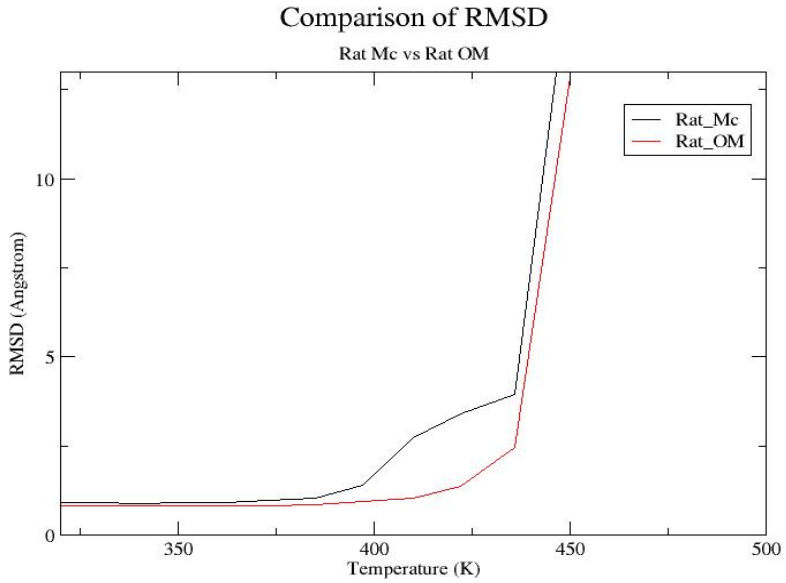


Figure 4.2 : Temperature vs RMS Deviations of rat Mc and rat OM protein

Table 4.2: RMSD of backbone atoms from starting structure

Temperature (K)	Rat_Mc (Holo protein) Å	Rat_OM (Holo protein) Å
320	0.92	0.83
330	0.90	0.84
340	0.90	0.83
351	0.90	0.83
362	0.93	0.81
373	0.96	0.82
385	1.05	0.86
397	1.40	0.94
410	2.72	1.03
422	3.41	1.38
436	3.94	2.47
450	16.07	12.84

4.3 Overall Radius of Gyration

The radius of gyration of the entire protein in the temperature range 320 K – 450 K, was in the range of 12.57 – 12.98 for the Microsomal protein and 12.59 – 12.87 for the OM protein. The overall radius of gyration values were very similar for both proteins and were comparable to the values obtained from previous Molecular

Dynamics (Mc form – 12.5 and OM form – 12.4).[22] From the Table 4.3, we observe that the increase in the radius of gyration is steeper in the Microsomal form than the OM form, and significant increase in the Rgy values are observed only after 410 K in case of the OM protein.

Table 4.3: Average Radius of Gyration (Entire protein)

Temperature (K)	Rat_Mc (Holo protein)	Rat_OM (Holo protein)
320	12.57	12.59
330	12.60	12.61
340	12.66	12.63
351	12.70	12.66
362	12.77	12.70
373	12.84	12.77
385	12.98	12.87
397	13.35	12.99
410	14.33	13.13
422	14.99	13.43
436	15.79	14.17
450	26.34	22.03

4.4 RMSD and Energy Histograms

From replica exchange results we obtain several conformations and in order to observe the distribution of energies and RMSD values of the conformations we construct histograms. A histogram summarizes the distribution of all data points within a particular range or class interval. Hence generally in the X-axis we have the class interval or the range within which data points lie and Y-Axis represents the number count or frequency of times the data points were found within the class interval.

The RMSD values were tabulated using the rexinfo.pl utility from the MMTSB toolset. Using this perl script we can obtain all the conformation PDB files at a given temperature and using these RMSD and Energy Histograms are constructed. The RMSD calculated over here varies from the RMSD value calculated previously from average trajectory structure and this might be due to difference in Orientation. In this histogram we observe the range in within all conformations of cytochrome b5 deviate compared to starting structure. Both the rat Microsomal and Outer Mitochondrial histograms are shown in the following figure. In the Microsomal protein, it is found that most of the conformations have a RMS deviation of about 4 Angstroms and as the temperature increases the RMSD values increase to about 5.1 Angstroms.

The RMS Deviations of the OM protein were about 1.4 Angstroms in the temperature region (320 K – 340 K) and as the temperature increased, (351 K – 385 K), the RMSD values gradually increased to about 2 Angstroms. There was a drastic increase in the RMSD values to 3.5 Angstroms, for temperature region above 400 K. Overall the OM proteins had lower RMS Deviations as compared to the Mc proteins. The following histogram, lists the RMSD profiles for temperatures below 400 K.

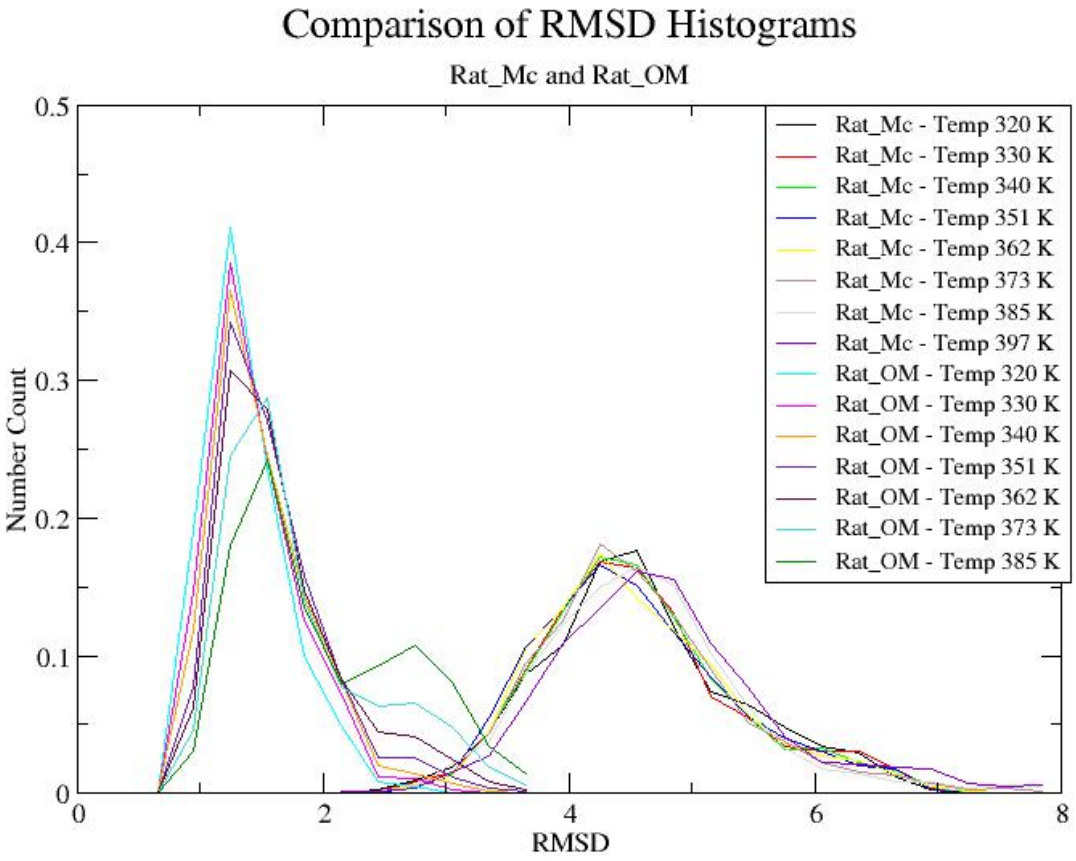


Figure 4.3: RMSD Histograms of Rat Mc and OM proteins at different temperatures

4.5 Energy Histograms

The Energy Histogram of the Microsomal protein ranged between -2900 Kcal/mol to –2000 Kcal/mol, whereas the OM protein histogram ranged between –2800 to –1980 Kcal/mol. The energy histogram of Microsomal protein was broader than the OM energy histogram. In order to identify the temperature range to conduct the simulation and to identify optimal number of replicas to be used , we conducted a trial run. During this trial the main criteria was that the swap acceptance probability was between 0.2 –0.3. This interval is a standard which has been arrived at by several other scientists.

During the trial run it was observed that, there was considerable overlap in the neighboring energy profiles and the swap acceptance probability was maintained between 0.2 –0.3 in case of Microsomal protein and was between 0.20-0.35. The histograms are represented as a function of Number Count and Energy values.

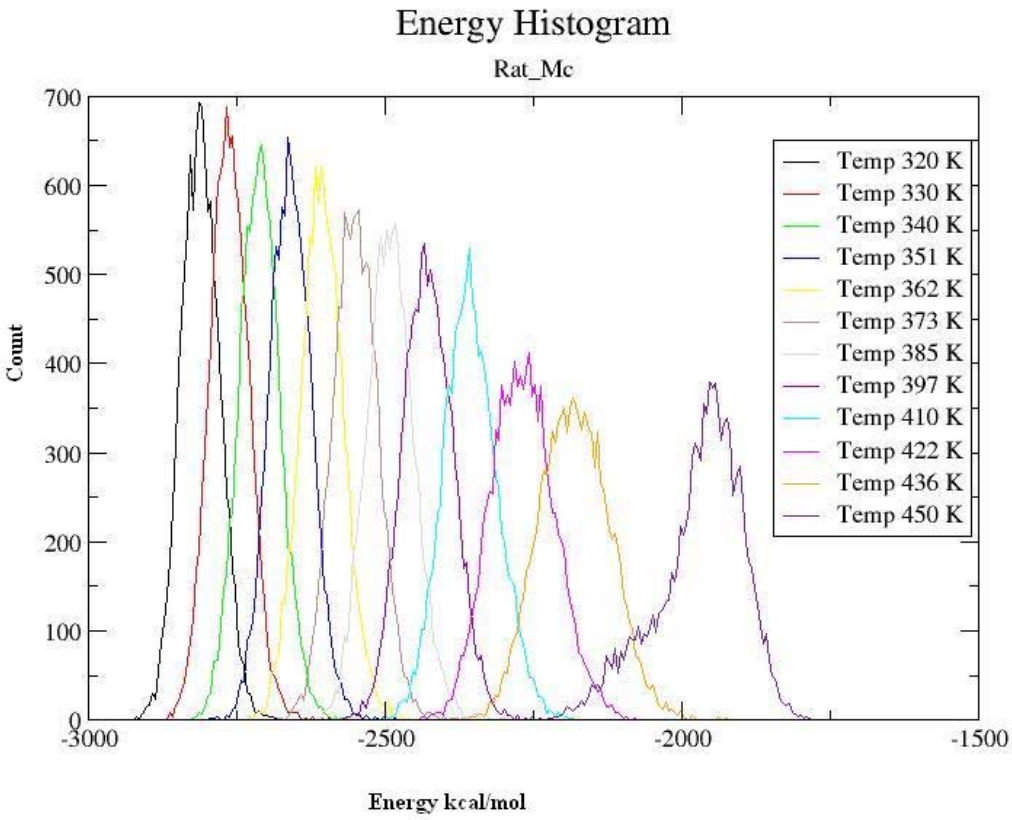


Figure 4.5: Histogram of energy profiles – Rat Microsomal protein

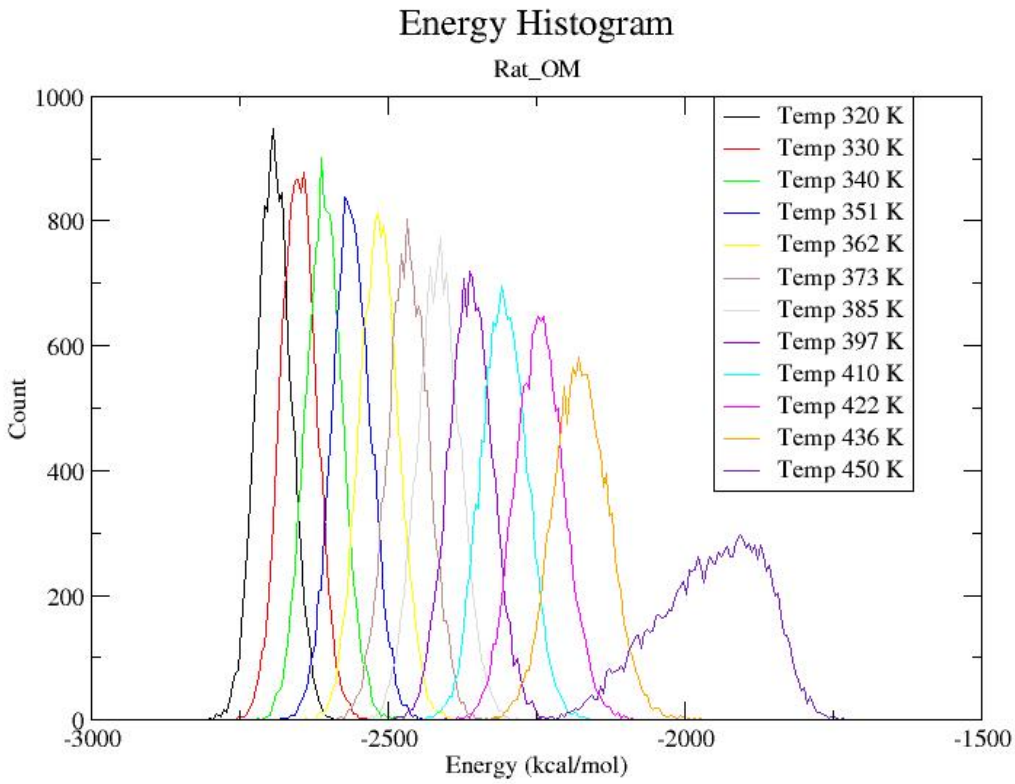


Figure 4.6: Energy Histogram of Rat OM protein.

4.6 Comparison of Core 1 and Core 2

Previous studies have shown that core 1 or the heme-binding core contributes to the functional role of cytochrome b5 and that core 2 contributes towards the structural integrity of the protein. The RMS Deviations, solvent accessibility and core sizes have been calculated and graphs comparing the rat microsomal and rat outer mitochondrial proteins have been drawn.

A comparison of the core 1 RMS Deviations shows that, in the temperature region 320 K -397 K there is no significant difference between the two proteins and the RMSD values are around 0.9 Å. When the temperature is above 410 K, the slope of the RMSD curve for the Mc protein is steeper than that of OM protein, suggesting that some degree of thermal unfolding. Similar to the observation in core 1, the RMS Deviations of core 2 are quite similar until about 397 K. At temperatures above 397 K, the core 2 RMS deviations are higher for Microsomal protein than for OM protein. From Figure 4.7, we also observe that there is a steep increase in the RMSD at 410 K in case of the Microsomal protein and a similar increase in the slope is observed at 422 K in case of the OM protein. As core 2 contributes to the structural integrity of the protein, this is consistent with the stability results, which indicate that OM form is more stable than the Microsomal protein.[45]

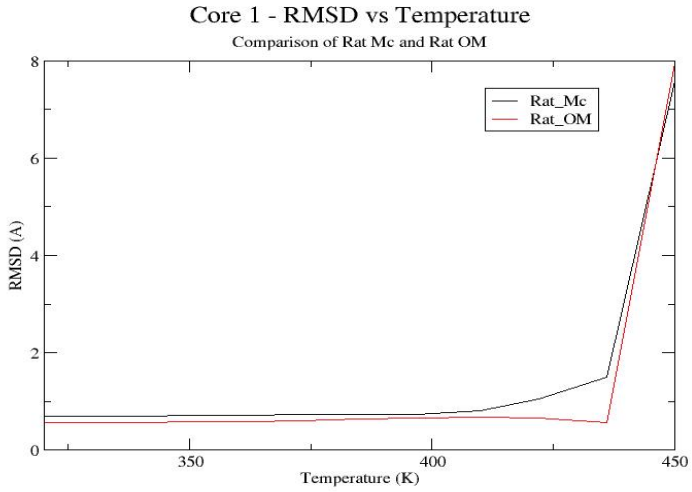


Figure 4.7: RMS Deviations of Core 1 (32:76) from starting structure.

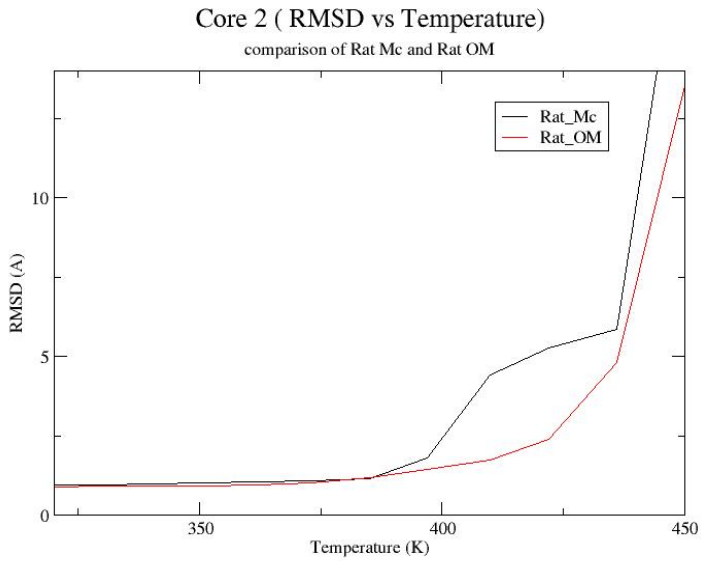


Figure 4.8: RMS Deviations vs. Temperature for Core 2

The overall radius of gyration, for both the Microsomal and Outer Mitochondrial protein were around the same range in the temperature region (320 K – 397K), accordingly the Core 1 radius of gyration is about 10.8 for both the proteins. When the temperature is about 410 K, the radius of gyration increases to 11.2 for the Microsomal protein. We observe a similar trend in the RMSD values, suggesting the structural deformation of the protein. Unlike the Microsomal protein, which has a gradual increase in the core 1 size, the OM protein has a drastic increase in the Radius of gyration when the temperature reaches 450 K. The Comparison of Core 2 radius of gyration suggests show that core 2 is more labile than core 1, reflecting presence of enforced Fe-His bonds in core 1.

Similarly from melting curve we also observe a bump in the graph, suggesting a presence of intermediate in the unfolding path of Microsomal holo protein.

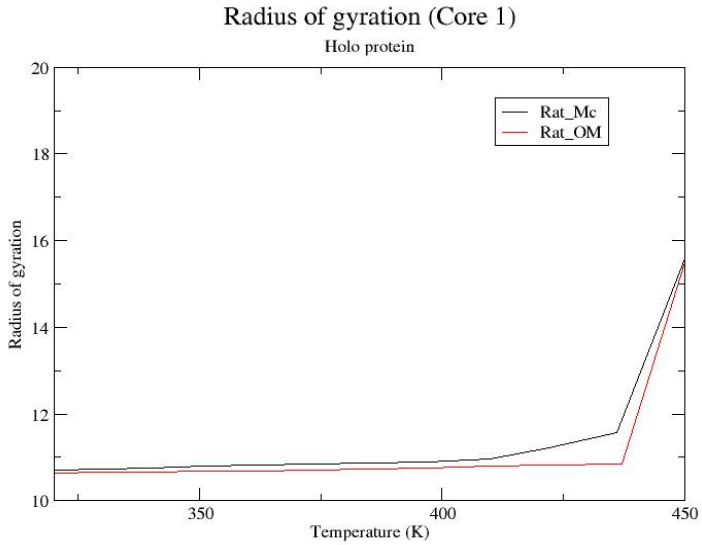


Figure 4.9: Radius of gyration of core 1 as a function of Temperature

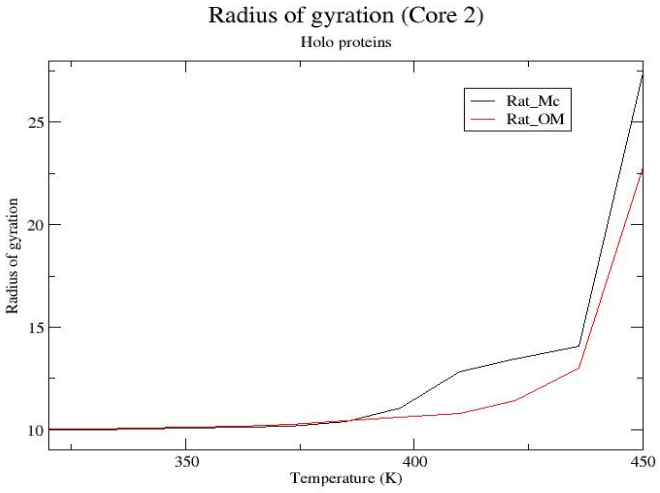


Figure 4.10: Radius of gyration of core 2 as a function of Temperature

The following graph shows the solvent accessible surface area (core 1) of rat OM and rat Mc at a temperature of 320 K. Core 1 of the OM protein was more exposed to the solvent than the core 1 of microsomal protein. At 320 K the average SASA value of rat Mc was 2892 ± 68 $Å^2$, whereas the SASA average value of rat OM was about 3201 ± 69 $Å^2$. However the Solvent Accessibility of the core 2 is around 2500 $Å^2$ for both the proteins. The SASA profiles of Core 1 at different temperatures are depicted as a histogram and in the temperature range 320 – 373 K, uniformly we observe that the solvent accessibility is higher in case of the rat OM protein.

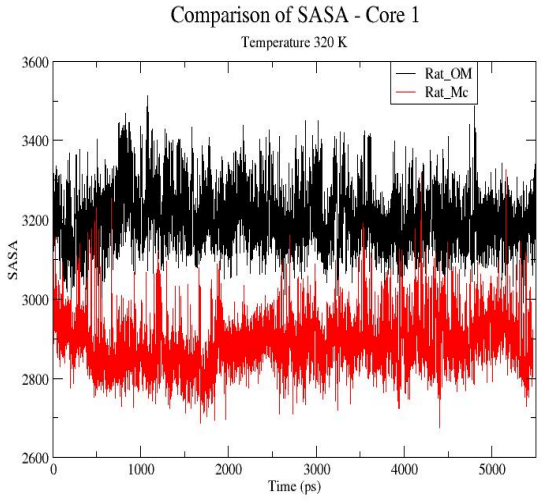


Figure 4.11: SASA of core 1 at 320 K

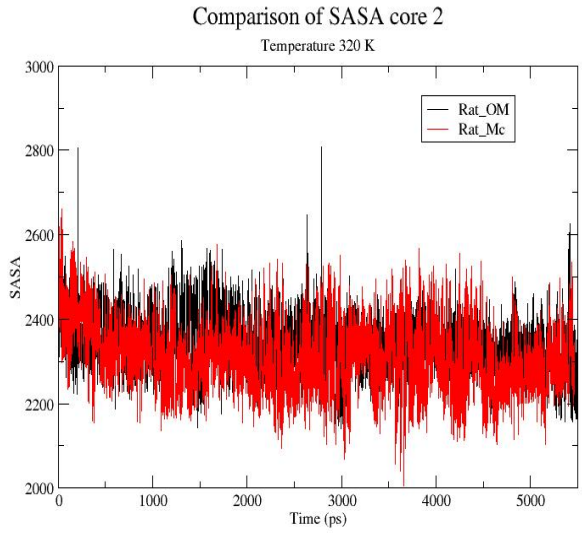


Figure 4.12: SASA of core 2 at 320 K

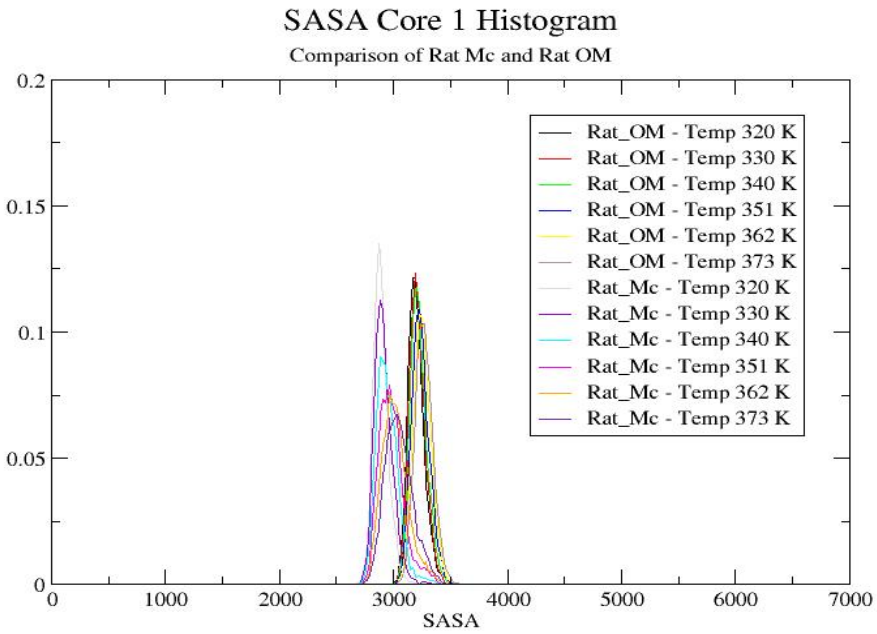


Figure 4.13: Histogram of SASA values of Core 1 at different temperatures.

4.7 SASA of the entire protein:

The solvent accessible surface area of Rat Microsomal protein at 320 K was 5200.78 $Å^2$ and the SASA value of rat OM cytochrome b_5 at 320 K was 5537 $Å^2$. As observed from the Core 1 solvent accessibility values, the SASA values of the Rat OM protein were higher than the Rat Mc protein. In case of both the proteins there was a gradual increase in the solvent exposure as the temperature increased. Though initially the solvent accessibility of rat Mc protein was lower than that of OM protein, at high temperatures above 397 K, the Mc form was more exposed to solvent , this might be due to de-stabilization of the helices.

Table 4.4 : SASA of Rat Mc and Rat OM protein, residues (4-85)

Temperature (K)	Rat_Mc (Holo protein) $Å^2$	Rat_OM (Holo protein) $Å^2$
320	5200.78	5537.28
330	5247.29	5561.95
340	5306.36	5592.47
351	5359.79	5631.38
362	5416.79	5677.10
373	5486.43	5740.61
385	5581.07	5822.71
397	5766.37	5935.61

410	6189.92	6056.64
422	6647.49	6261.78
436	7147.44	6682.41
450	10716.6	9997.49

Solvent accessible surface area of heme residue was calculated for both Mc rat cytochrome b_5 and OM rat cytochrome b_5. It was observed that at 320 K, the heme residue in OM cytochrome b5 had a higher exposure to solvent (SASA 260.6 $Å^2$) and Mc cytochrome b5 had a SASA value of 256.68 $Å^2$. The melting curve of heme SASA value is shown in Figure 4.15. As observed from the melting curve the difference in solvent exposure of heme is very less to be significant statistically.

Table 4.5 : SASA of Heme - Rat Mc and Rat OM protein.

Temperature (K)	Rat_Mc (Holo protein) $Å^2$	Rat_OM (Holo protein) $Å^2$
320	256.68	260.62
330	257.56	263.57
340	258.03	265.82
351	261.09	269.86
362	263.84	274.56
373	266.86	279.77

385	269.63	284.92
397	270.31	289.74
410	269.65	291.54
422	288.90	294.31
436	317.77	291.64

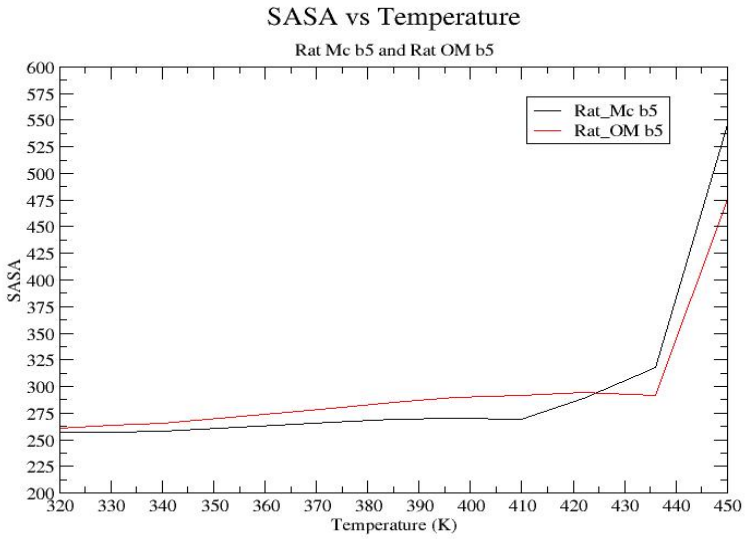


Fig 4.14: SASA of heme residue vs. Temperature

4.8 Native Contacts

The ρ value represents the amount of native contact as compared to the reference structure. The criterion for selecting residues in contact is that the inter-residue distance must be less than 4.2 Angstroms. The fraction of native contact i.e ρ value, is calculated using a sigmoidal function, which takes into account even residue pairs that are slightly apart.The residue contact maps were calculated with respect to the starting structure using the Perl utility contact.pl. At 320 K, the Rat Mc protein had ρ value of 0.8758 and the Rat OM protein had a value of 0.9150. As the temperature increased, some of the residue contacts were lost, and significant reduction in ρ values were observed only at temperatures above 397 K. At 450 K the rat Mc had a ρ value of 0.6722 and the OM protein had a ρ value of 0.7783. As compared to the OM protein, fraction of native contact was lower for the Microsomal protein even at a temperature of 320 K. Figure 4.15 compares the fraction of native contact (ρ value) of both the proteins at different temperatures. The histogram of ρ values for rat Mc protein broadens as it reaches very high temperatures unlike the OM protein. This broadening shows that more and more folded structures are being converted to unfolded structures, which is reflected by loss of native contacts. The ρ values obtained along with structural analysis results support previous inference that the melting temperature of OM protein was higher as compared to the Mc protein.

Table 4.5: Native residue contacts as compared to starting structure.

Temperature	Average Rho – Rat Mc	Average Rho – Rat OM
320 K	0.88	0.92
330 K	0.88	0.91
340 K	0.88	0.91
351 K	0.88	0.91
362 K	0.88	0.91
373 K	0.87	0.91
385 K	0.87	0.91
397 K	0.87	0.91
410 K	0.87	0.90
422 K	0.87	0.90
436 K	0.85	0.90
450 K	0.67	0.78

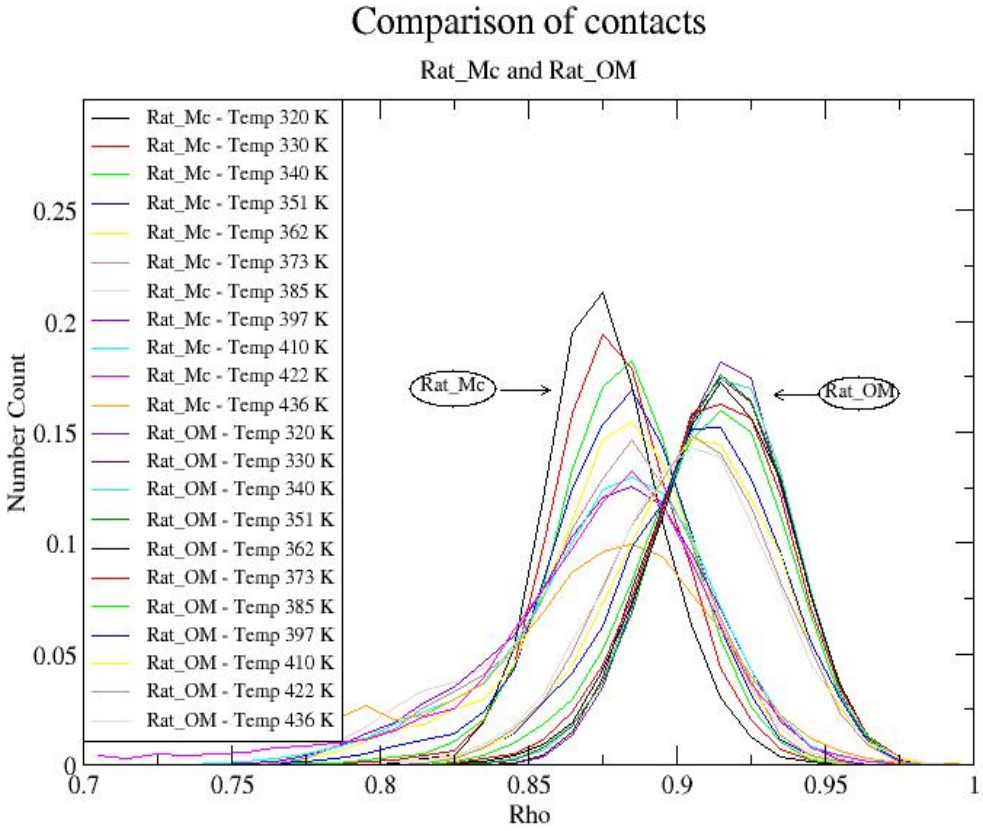


Fig 4.15: Histogram of fraction of native contact

4.9 Secondary Structure Analysis

The hydrogen bond N…O distances were calculated for both the proteins at different temperatures. At every temperature, the percentage of number of hydrogen bonds to total number of bonds were calculated. The cut off N…O distance was given as 3.6 Angstroms. At 320 K, all the four helices were well maintained in case of both the rat Mc b_5 and OM protein b_5 protein. Helix α1 consisted of residues 9:13. At the lowest temperature this helix retained about 98 % helicity in the rat microsomal b5 and retained about 99 % in the rat OM b5 protein.

The rat microsomal b5 protein lost hydrogen bonds at a faster rate as compared to OM b5 protein. This is reflected by the data that rat OM protein retained about 97 % of its hydrogen bonds at 385, whereas at this temperature the rat Mc protein had only 87 % of its hydrogen bonds. While comparing the differences between rat microsomal b5 and rat OM b5 it was found that the helix α1 was more well structured in rat OM b5 protein. Since the helix α1 lies in the core 2, this is consistent with experimental results that core 2 is more robust in the OM b5 protein as compared to Mc b5.

Table 4.6: Hydrogen bond percentage vs. Temperature

Temperature	Rat Mc **Helix α1** 13 N 9 O	Rat Mc **Helix α1** 14 N 10 O	Rat OM Helix **α1** 13 N 9 O	Rat OM **Helix α1** 14 N 10 O
320 K	98.70 %	87.55 %	99.59 %	75.88 %
330 K	98.40 %	84.20%	99.49 %	74.60 %
340 K	98.00 %	83.30 %	99.52 %	74.68 %
351 K	96.60 %	83.35 %	99.30 %	73.80 %
362 K	93.60 %	84.30 %	98.88 %	72.61 %
373 K	91.10 %	85.50 %	97.90 %	70.70 %
385 K	87.80 %	85.99 %	97.10 %	69.50 %
397 K	84.80 %	82.88 %	94.70 %	68.60 %
410 K	76.56 %	72.70 %	91.1 %	67.70 %
422 K	70.30 %	63.10 %	81.80 %	65.80 %
436 K	62.10 %	58.60 %	81.50 %	63.30 %
450 K	7.00 %	7.60 %	16.50 %	10.60 %

The helix α3 consisted of residues 44:49 and it was well maintained at lower temperatures. The N…O bond distance between residues 45 and 49 was greater than 3.6 Angstroms for most part of the trajectory. But the 50 N…46 O bond distances were within 3.6 Angstroms and overall the helix α3 was well maintained.

With temperature increase there was considerable loss in the secondary structure and incase of rat Mc protein from about 97 % at 320 K, the hydrogen bond percentage was 74 % at 385 K. Whereas the hydrogen bond percentage of rat OM protein, decreased from 95 % (320 K) to 87 % (385 K). Hence when compared to other helices this helix α3 exhibited more loss of secondary structure.

Table 4.7 : Helix α3 hydrogen bond percentage

Temperature	Rat_Mc **Helix α3** 48 N 44 O	Rat_Mc **Helix α3** 49 N 45 O	Rat_OM **Helix α3** 48 N 44 O	Rat_OM **Helix α3** 49 N 45 O
320 K	96.90 %	54.40 %	95.30 %	60.48 %
330 K	95.40 %	52.60 %	94.60 %	60.27 %
340 K	93.20 %	50.10 %	93.55 %	58.90 %
351 K	90.40 %	47.80 %	92.40 %	57.48 %
362 K	85.80 %	47.11 %	91.90 %	56.59 %
373 K	80.50 %	45.05 %	89.60 %	54.20 %
385 K	74.10 %	42.90 %	87.10 %	55.12 %
397 K	72.80 %	41.30 %	84.30 %	55.65 %
410 K	73.60 %	39.09 %	81.10 %	55.31 %
422 K	61.50 %	36.09 %	76.80 %	53.90 %
436 K	54.24 %	32.93 %	72.60 %	52.08 %
450 K	43.30 %	39.09 %	44.00 %	38.69 %

Helix α4 consists of residues 55 to 61 and had a H-bond percentage of 85 % for rat Mc protein and a percentage of 81 % for rat OM protein, at 320 K. Similar to other helix hydrogen bond distances, percentage deviation was higher in case of the rat Mc protein than the rat OM protein. This helix was extremely stable in case of the rat OM protein and retained about 97 % of its hydrogen bonds at 422 K. With respect to helix α3, with temperature increase not much difference was between isoforms.

Table 4.8 : Helix 3 – Hydrogen bond percentages

Temperature	Rat Mc **Helix α4** 59 N 55 O	Rat Mc **Helix α4** 60 N 56 O	Rat OM **Helix α4** 59 N 55 O	Rat OM **Helix α4** 60 N 56 O
320 K	85.12 %	99.50 %	80.70 %	99.60 %
330 K	84.50 %	99.30 %	79.08 %	99.27 %
340 K	82.90 %	99.08 %	77.87 %	99.15 %
351 K	83.98 %	98.88 %	76.70 %	99.17 %
362 K	82.48 %	98.45 %	74.80 %	98.80 %
373 K	81.63 %	98.25 %	74.68 %	98.23 %
385 K	81.50 %	97.37 %	74.62 %	97.96 %
397 K	81.47 %	96.48 %	74.71 %	97.82 %
410 K	79.90 %	94.10 %	72.50 %	97.40 %
422 K	73.57 %	87.10 %	69.12 %	96.70 %
436 K	64.90 %	72.30 %	67.10 %	94.60 %
450 K	19.70 %	23.00 %	41.80 %	55.10 %

The helix α5 consists of residues 65:69 and of all the helices; this helix was very stable in both the rat Mc and rat OM protein. A t 320 K, the rat Mc protein retained about 97.9 % of the hydrogen bonds and in rat OM protein it had about 99.69 % of hydrogen bonds. With temperature increase the loss of helicity was immediate in case of Microsomal b5 protein whereas the rat OM b5 protein retained most of its hydrogen bonds with temperature increase. Hence this helix α5 is more stable in case of the rat OM b5 protein as compared to Microsomal protein. This supplements experimental results, which implicate Leu 71 residue to play an important role in stabilizing the rat OM b5 protein.[44]

Table 4.9: Helix α5 – Hydrogen bond distances as percentage

Temperature	Rat Mc **Helix α5** 69 N 65 O	Rat Mc **Helix α5** 70 N 66 O	Rat OM **Helix α5** 69 N 65 O	Rat OM **Helix α5** 70 N 66 O
320 K	98.60 %	97.90 %	99.69 %	98.57 %
330 K	97.14 %	95.70 %	99.57 %	98.23 %
340 K	94.80 %	92.80 %	99.43 %	97.70 %
351 K	92.30 %	89.70 %	99.08 %	96.86 %
362 K	91.01 %	87.40 %	99.02 %	95.80 %
373 K	88.35 %	84.20 %	98.48 %	94.70 %
385 K	85.70 %	81.10 %	97.66 %	93.53 %
397 K	82.10 %	77.00 %	96.63 %	91.69 %

410 K	74.80 %	72.70 %	94.50 %	90.30 %
422 K	61.80 %	63.40 %	90.80 %	85.90 %
436 K	40.1 %	43.90 %	87.01 %	81.70 %
450 K	26.00 %	28.00 5	26.84 %	24.40 %

4.10 Comparison of structural fluctuations in Apo and Holo proteins.

The structural fluctuation of rat microsomal b5 protein at 320 K is compared to the apo protein results and we observe that the main difference in the RMS Fluctuations occurs in the core 1 region. In the core 2 regions, the fluctuations observed in the apo protein are those, which are also observed in the holo protein. Hence this suggests that the absence of heme group affects the core 1 residues mainly and the implications are localized to the core 1 (32:76) region. In the holo microsomal b5 protein we observe increased mobility in residue number 18, which is not observed in the apo microsomal b5 protein.

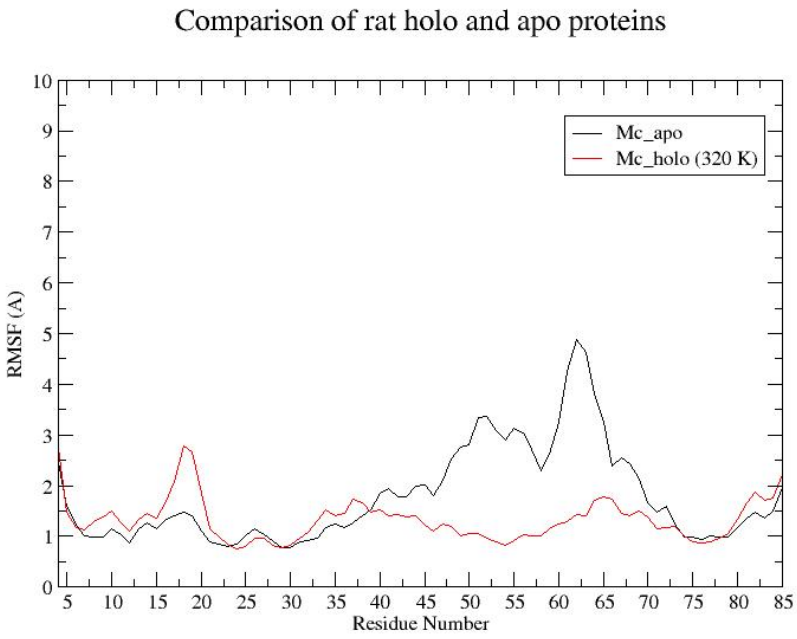


Fig 4.16: RMSF vs Residue Number – Microsomal Holo cytochrome b5 and Microsomal Apocytochrome b5

A similar comparison of apo OM cytochrome b5 simulations and holo OM cytochrome b5 simulations show that the contours of both the graphs are very similar. This suggests that even in the absence of heme group the heme pocket is maintained more intact and indicates presence of hydrophobic interactions. However since our simulations of apo OM b5 protein was based on a model we cant concretely arrive at conclusions.

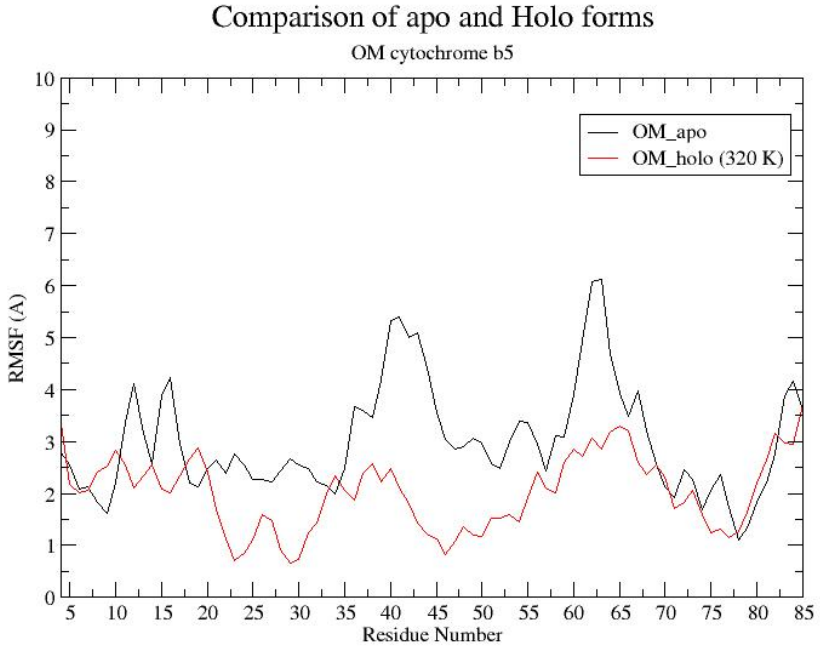


Fig 4.17 – RMSF vs Residue Number for OM holo cytochrome b5 and OM Apocytochrome b5.

4.11 DISCUSSION

The Replica Exchange method was used to conduct 4.4 ns simulations of the rat Microsomal and Outer Mitochondrial proteins. The conformational analysis of the results is conducted and the temperature variations of structural fluctuations, solvent accessibility, and secondary structure have been reported. The RMS Deviation from the starting structures was calculated and the Microsomal protein had a larger deviation as compared to the rat OM protein. However the RMS Fluctuation of the rat OM protein was slightly larger as compared to the Microsomal protein.

Since we have employed the replica exchange method to conduct the dynamics, the atomic fluctuations of individual residues can be compared with hydrogen exchange experiment results.[42] In the Microsomal protein the region between 16-20 , the end of the helix 2 (36 – 45) and the histidine 63 loop region (61-66) were all highly mobile. When compared to the experimental results, all these regions corresponded to regions with faster exchanging rates. Similarly in the rat OM protein the regions, which exhibited high RMS fluctuations, were 35-41 and the region between 60 – 65. The comparison of RMS Deviations and Fluctuations show that the microsomal form exhibited higher deviations and was comparatively more mobile than the OM protein. Individual residues like 27, 37 and 73 were quite flexible in the OM protein, whereas these regions exhibited low atomic fluctuations in the rat Mc protein. Though some of the regions corresponded to the Hydrogen exchange results, in during our simulations we couldn't find significant difference between the overall flexibility of the proteins.

Melting curves of RMSD vs. temperature, RMSF vs. Temperature were plotted for both Microsomal and OM holo proteins. With the increase in temperature the rat Mc protein exhibited higher RMSF values, whereas the Rat OM protein was quite stable till 397 K. The increased RMS Deviations and RMS Fluctuations suggest that the Microsomal protein undergoes unfolding due to temperature. This is in line with our expectations as the OM holo protein has a higher stability according to chemical denaturation and thermal denaturation experiments.[45] Also the melting curves of the Microsomal cytochrome b_5 showed a bump at around 400 k, which supports the theory that the Microsomal protein unfolds in two stages unlike the rat OM protein where no such bump was observed.

From the RMSD histograms at higher temperatures above 362 K, the Mc histograms broaden into a Gaussian like distribution indicating the existence of unfolded conformations. Similarly the tail region of the OM distribution broadens, thereby suggesting conversion of folded conformations into unfolded structures.

The RMS Deviations of the heme-binding core (core 1) with respect to temperature is observed and we find that the core 1 of rat OM protein is more stable for a longer temperature range. Whereas in case of the core 2 of Microsomal proteins we observe a steep increase in the RMSD values as temperature increases and in the OM protein, we find a gradual increase in the RMSF. Therefore the core 1 in the rat OM protein is

more resistant to thermal denaturation than the core 1 of rat Mc protein. A comparison of core 1 and core 2 melting curves also show that core 2 is more mobile than core 1 reflecting the presence of enforced Fe-His bonds. Also in both microsomal b5 and OM b5 we observe that the RMSF, RMSD and radius of gyration values differ significantly in core 2 at around 385 K, whereas such a difference in values is not observed in core 1. This support experimental results, which observe that core 2, unfolds prior to core 1.[47]

The secondary structure analysis of rat microsomal b5 protein showed that the helices α1(9-13) and α4 (55-61) retained its helicity for most part of the simulation. The α3(44-49) helix was also well maintained but it lost some hydrogen bonds towards the end of the helix. Whereas in case of rat OM b5 protein the helices α1(9-13), α4(55-61) and α5(65-71) were stable and showed high helical content. The core 2 region of rat OM protein was more robust as compared to the core 2 region of rat Mc b5 protein. Also it was observed that with increasing temperature the helices in core 2 region showed higher loss of secondary structure supporting the results that core 2 unfolds prior to core 1 region.

The Solvent accessibility of both the core 1 and core 2 are calculated at different temperatures. At 320 K, the solvent accessibility of core 1 in the rat OM protein is higher than the rat Mc protein, whereas no difference in SASA value is observed with respect to core 2. The solvent accessible area for heme residue was calculated and

heme in rat OM b5 had a higher solvent exposure as compared to rat Mc b5. This is also consistent with experimental results, which show that heme in OM cytochrome b5 is more exposed to solvent than microsomal b5 protein.[46]

Apart from the Secondary structure analysis, the number of native contacts remaining as compared to the starting structure was also calculated for both the proteins. The rat OM had a higher amount of residual native contact as expressed by higher ρ value. Hence the rat OM protein was more structured at temperatures between 340 – 397 K.

4.12 CONCLUSIONS

The replica exchange method was used to simulate structural fluctuations of holo proteins to serve dual purpose of increased conformational sampling and to study temperature variation of several molecular properties. The RMS deviations of the rat microsomal protein from starting structure were higher than the rat OM protein. With increasing temperature the fluctuations in rat microsomal protein increased steeply whereas the rat OM protein was mobile only at very high temperatures. The plots of RMSF vs. Residue number at 320 K, showed that in both the isoforms of cytochrome b5, the Histidine loops (36-40) and (60-65) exhibited larger RMS Fluctuations along with the (16-20) region. A comparison of the RMS Fluctuations of all residues also showed some notable differences between the holo proteins. The β2, β3 and β 4 strands were conformationally flexible in the rat Microsomal proteins, whereas they

had lower RMSF values in the rat OM protein simulation. The variation of structural properties with respect to temperature showed that the rat OM protein was more stable than the Mc form and had a higher melting temperature. At higher temperatures the holo proteins unfolded and this was observed in the RMSD and Energy Histograms, as population of folded structures decreased and widening of histogram occurred. However complete unfolding might not be observed due to presence of restraining Fe-His bonds.

A comparison of the core 1 and core 2 RMSF, RMSD and Radius of gyration, showed that any changes in the properties in core 1 was associated with concurrent change in the core 2, thereby suggesting a cooperative unfolding process due to thermal denaturation. With increasing temperature increased mobility and radius of gyration was observed in core 2, however such an increase in values were observed much later in core 1, reflecting that structural disintegration of core 2 before core 1. Another observation from melting curves was the presence of an intermediate, suggesting biphasic unfolding in case of rat microsomal b5 protein, which wasn't observed in rat OM form.

Using secondary structure results we compared helicity in core 1 and core 2, and they showed that core 1 retained its helicity whereas core 2 lost hydrogen bonds, supporting our results which suggested core 2 might unfold prior to core 1. Also another difference was that core 2 region of rat OM b5 protein was more robust than rat Mc b5 protein. Apart from this the helix α5 was well structured and retained helicity in rat OM b5 protein, whereas in the rat Mc b5 it was dis-structured. The solvent accessibility of core 1 in the rat OM form was higher than the rat Microsomal form. As the temperature increased the number of native contacts reduced in the proteins and percentage of hydrogen bonds present also decreased gradually.

To summarize our simulations show that the rat OM protein exhibited lower fluctuations, higher structural integrity, and resistance to thermal unfolding due to extremely stable core 1 as opposed to Microsomal protein. Hence using replica exchange we were able to observe detailed fluctuations of individual residues and understand the effects of temperature increase on the overall stability, structural integrity and dynamics of the holo proteins.

5. MODELING PROTEIN LOOPS

Homology modeling or comparative modeling refers to building a model structure for a unknown protein from its sequence based on known structures, also known as templates.[47]This is based on the premise that proteins with similar sequences have similar structures. The modeling of protein loops can be achieved using various techniques like modeling by assembly of rigid bodies, modeling based on spatial restraints, modeling by segment matching etc. [48]Though there are several algorithms, most of these work well in case of short loop regions. Another major limitation is prediction of unknown structure, which is not structurally divergent to all known structures; in that case comparative modeling becomes difficult. Hence using constrained minimization and replica exchange method we tried to reconstruct long segment of proteins. The loop region selected for reconstruction was obtained from CASP 4 modeling targets.[47] These loop regions were structurally divergent when compared to determined PDB structures.

5.1 METHOD

All the regions, which were reconstructed, are shown in Table 5.1 and about six loop segments with varying length (15-34) were chosen. Throughout reconstruction of the loops, the CHARMM 19 parameters were used along with the Effective Energy Function 1 (implicit solvent model).[33]

This model uses the CHARMM 19 polar hydrogen energy function along with an excluded volume solvation model.[33] The Replica Exchange method was used to increase conformational sampling.

Table 5.1: List of loop segments, which were reconstructed

Loops	PDB file	Residues	Length
T090	1G9Q	77-91	15
T109	1J9A	48-81	34
T113	1E6W	203-223	21
T114	1GH5	51-55	15
T123	1EXS	65-82	14
T125	1GAK	94-118	25

Step 1: The initial coordinate of the entire protein is built using the sequence of residues and reference starting structure. The rest of the protein is deleted except the loop region and stem of the loop.

Step 2: The initial model of the loop is constructed in this step. The loop is reconstructed using a series of constrained minimization steps. The parameters required for the calculation of solvation free energy G_{solv} is given as input through a parameter file. The loop region coordinates are rebuilt with the exception of stem1

region. Then minimization is performed gradually by slowly increasing the force constant of the harmonic constraint placed. Initially the minimization is conducted using the 100 steps of Steepest Descent method, followed by Langevin dynamics. The minimization is performed iteratively with increasing harmonic force constant value. The RMS Deviation from starting structure is calculated periodically.

Step 3: In this step, the loop region, which has been modeled, is attached to the protein. Initially the entire protein coordinates are generated using the PDB file and residue sequence. Using that all the residues in loop region are overwritten, thereby now our model loop is incorporated into the entire protein. Then the protein is fixed and the loop region is minimized once again using ABNR and Powell algorithms.

Step 4: Replica Exchange simulations were conducted using the EEF1 parameter and about 12 simultaneous simulations were conducted in the temperature range 280-1000 K.The length of the replica exchange is about 7.5 ns. The protein is fixed and therefore several different conformations of the loop region is generated through replica exchange.

Step 5: The trajectory obtained from REMD is again minimized using ABNR and Powell algorithms. The output file records the energy , RMS Deviations of all conformations at different temperatures.

Step 6: Picking the best possible conformations based on a scoring function. In our case we tried picking the best conformations based on the lowest energy criteria. The lowest energy was chosen and corresponding RMSD value was ranked. Similarly since the best confirmation would have the least amount of deviation from starting structure, we chose the conformation with the lowest RMS Deviation and ranked the corresponding energy.

5.2 RESULTS

The RMSD-L values correspond to the RMSD values obtained by Rohl et al. using the Rosetta algorithm. In case of the t090 , t114 and t123 loops the minimum RMSD values were comparable to the deviations observed using the Rosetta method. All these three loop regions were about 15 residues in length , hence in these cases our method of prediction resulted in conformations with lower RMSD.

Table 5.1: Conformation ranking based on structure with lowest RMSD

Loops	**Min RMSD**	**Energy** (corresp min RMSD)	**Rank** (min rmsd)	**RMSD-L**
T090	2.635	-1072.69	3894	3.54
T109	5.847	-1779.0	9340	4.00
T113	4.574	-1384.69	5568	2.91
T114	1.321	-1069.43	9484	2.22
T123	2.985	-1320.09	5535	3.88
T125	4.764	-1797.87	7844	0.84

.

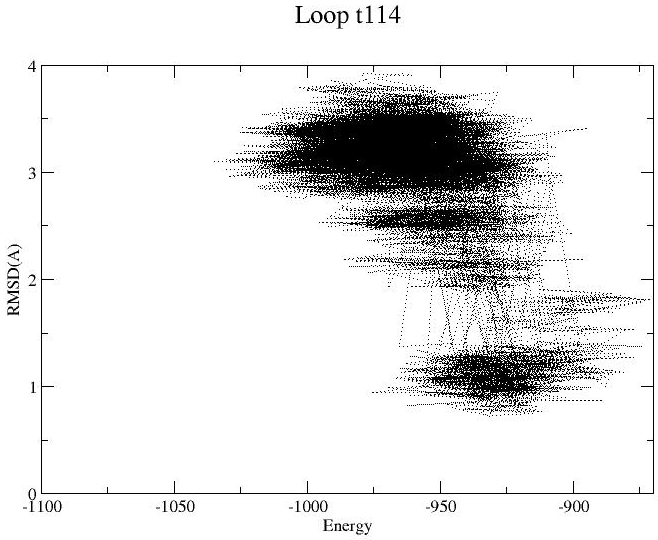
The Table 5.2 lists the ranking of conformations based on minimum energy. The RMSD values of all conformations with minimum energy are found out and are ranked in a set of all conformations. Even this kind of scoring method wasn't useful in finding the best possible conformation.

Table 5.2: Ranking of structures based on Minimum energy

Loops	**Min energy**	**RMSD** (corresp min energy)	**Rank** (w.r .t min energy)
T090	-1135.8	4.0116	9528
T109	-2315.31	7.9991	2960
T113	-1533.1	4.82595	3469

T114	-1160.18	3.55096	4961
T123	-1469.03	3.1336	132
T125	-2033.29	7.692	6719

The following plots show the variation of RMSD with respect to energy in two loops t114 and t123, in both the plots we find several conformations with low RMS Deviations and low energy value. However one of the challenges is to pick the optimized best possible conformation. The simulation results of t123 , showed a concentration of several structures which had low RMSD and energy values.



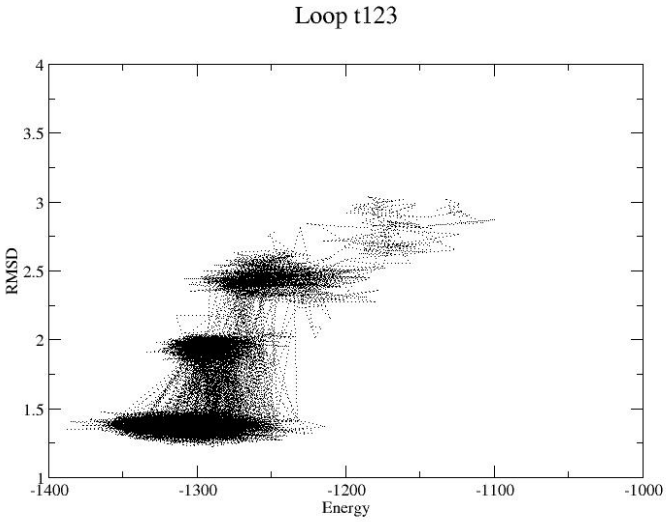


Figure 5.1: Plots of RMSD vs. Energy For t114 and t123 loop segments

5.3 CONCLUSIONS AND FUTURE WORK

One of the problems was the selection of best conformation based on a scoring function. Since we have several conformations, which have low RMSD values and low energy values, using a better scoring function would give us better results. The new scoring function could be based on both RMS Deviations and minimum energy. Also extending the replica exchange simulations would result in increased conformational sampling thereby enabling us to find the best possible conformation.

References

1. Leach, A.;(1991) A Survey of Methods for Searching through Conformational Space of Small and Medium-Sized Molecules. (Lipkowitz, K.B; Boyd, D.B.; Eds) *Reviews in Computational chemistry*, Vol 2, 4-10.

2. Vengadesan; Gautham, N.;(2005) A new conformational search technique and its applications, *Current Science*, 88(11), 1759-1770.

3. Mathews, .F.S.; Czerwinski, E.W. (1976) The Enzymes of biological membranes. (Martonosi, A, Ed.). New York: Plenum Press; 143-198.

4. Vergeres, .G. and Waskell, L. (1995) Cytochrome b5, its functions, structure and membrane topology. *Biochemie,* 77, 604-620.

5. Sanborn, R. and Williams, C.; (1950) The cytochrome system in the cecropia silkworm with special reference to the properties of a new component. *J.Gen.Physiol* ,33, 579-582.

6. Strittmatter, C.F.; Ball, E.G; (1954) The intracellular distribution of cytochrome components and of oxidative enzyme activity in rat liver. *J.Cell.Physiol*., 43(1), 57-78.

7. Hegesh, E.; Hegesh J.; Kaftory A.; (1986) Congenital methemoglobinemia with a deficiency in cytochrome b5. *N. Engl. J. Med.* 314 (12), 757-761.

8. Raw, I.; Mahler, H.R. (1959) Studies of electron transport enzymes. *J.Biol.Chem.* 234, 1867-1873

9. Porter, T.D. (2002) The roles of cytochrome b5 in cytochrome P450 reactions. *J.Biochem.Mol.Toxicol.* 16, 311-316.

10. Strittmatter, P.; Spatz, L.; Corcoran, D.; Rogers, M.J; Setlow B.; Redline R.; (1974) Purification and properties of rat liver microsomal stearyl coenzyme A desaturase. *Proc.Natl.Acad.Sci.* USA 71(11), 4565-4569.

11. Lederer, F., Ghrir, R., Guiard, B., Cortial, S., and Ito, A. (1983) Two homologous Cytochrome b5 in a single cell. *Eur.J.Biochem*., 132, 95-102.

13. Collins, F.S (2002) Generation and initial analysis of more than 15,000 full length human and mouse cDNA sequences. *Proc.Natl.Acad.Sci.* U.S.A. 99,16899-16903

14. Mitoma, J. and Ito, A. (1992). The carboxy terminal 10 amino acid residues of cytochrome b5 are necessary for its targeting to the endoplasmic reticulum.*EMBO J.* 11, 4197-4203.

15. Wang, L. , Cowley, A.B., Terzyan, S. , Zhang, X. and Benson, D.R (2007) Comparison of cytochromes b5 from insects and vertebrates. *Proteins: Structure Function and Bioinformatics* 67 (2) 293-304.

16. Parsons, D.F. and Williams, G.R. (1967). Isolation and purification of the outer membrane and inner membrane of the liver mitochondria. *Methods Enzymol.* 10, 443.

17. Storch E.M, Daggett .V., (1995). Molecular Dynamics of cytochrome b5: implications for protein-protein recognition. *Biochemistry*, 34(30), 9682-9693.

18. Ozols, J. (1989). Structure of cytochrome b5 and its topology in the microsomal membrane. *Biochim. Biophys. Acta* 997(1-2): 121-130.

19. Rodriguez-Maranon M.J, Qiu F., Stark R.E, White S.P., Zhang, X., Foundling, S.I., Rodriguez V., Schilling CL, Bunce, R.A, Rivera, M. (1996). 13C NMR spectroscopic and X-ray crystallographic study of the role played by mitochondrial cyochrome b5 heme propionates in the electrostatic binding to cytochrome c. *Biochemistry,* 35(50) 16378-16390.

20. Rivera M, Seetharaman R., Girdhar D., Wirtz M., Zhang X., Wang X, White S. (1998) The reduction potential of cytochrome b5 is modulated by its exposed heme edge. *Biochemistry*, 37(6), 1485-1494.

20. Silchenko, S., Sieppel, M.L, Kuchment, O., Benson D.R., Mauk, A.G., Altuve, A. and Rivera, M. (2000). Hemin is kinetically trapped in cytochrome b5 from rat outer mitochondrial membrane. *Biochem. Biophys. Res. Commun.,* 273, 467-472.

21. Altuve, A. Wang, L., Benson, D.R. and Rivera M. (2004) Mammalian Mitochondrial and microsomal cytochrome b5 exhibit divergent structural and biophysical characteristics.

22. Cowley, A.B, Rivera, M., Benson, D.R. Stabilizing roles of residual structure in empty heme binding pockets and unfolded states of microsomal and mitochondrial apocytochrome b5. (2004). *Protein Sci*., 13, 2316-2329.

23. Cheng, Q. , Benson, D.R, Rivera, M, Kuczera K., (2006) Influence of point mutations on the flexibility of cytochrome b5: Molecular Dynamics simulations of holo proteins. *Biopolymers*, 83, 297-312.

24. Storch, E. and Daggett, V. Consequences of heme removal: MD simulations of rat and bovine Apocytochrome b5 (1996). *Biochemistry*, 35(36), 11596 -11604.

25. Affentranger, R.; Tavernelli, I.; Di Iorio, E. E., (2006) A novel Hamiltonian replica exchange MD protocol to enhance protein conformational space sampling. *J. Chem. Theory Comput.* 2 (2), 217-228.

26. Sugita, Y.; Okamoto, Y., (1999) Replica-exchange molecular dynamics method for
protein folding. *Chem. Phys. Lett.* 314, (1-2), 141-151.

27. Fukunishi, H.; Watanabe, O.; Takada, S., (2002). On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.,* 116 (20), 9058-9067.

28. Jas, G. S.; Kuczera, K.,(2004). Equilibrium structure and folding of a helix-forming peptide: Circular dichroism measurements and replica-exchange molecular dynamics simulations. *Biophys. J.* **,** 87 (6), 3786-3798.
29. Falzone, C.J., Mayer, M.R., Whiteman, E.L., Moore, C.D., and Lecomte, J.T.J. (1996) Design challenges for hemoproteins: The solution structure of apocytochrome b_5. *Biochemistry,* 35, 6519–6526.

30. Moore, C.D. and Lecomte, J.T.J. (1990) Structural properties of apocytochrome b_5: Presence of a stable native core. *Biochemistry,* 29, 1984–1989.

31. Bhattacharya, S., Falzone, C.J., and Lecomte, J.T.J. (1999) Backbone dynamics of apocytochrome b_5 in its native, partially folded state. *Biochemistry*, 38, 2577–2589.

32. Manyusa, S. and Whitford, D. (1999) Defining folding and unfolding reactions of apocytochrome b_5 using equilibrium and kinetic fluorescence measurements. *Biochemistry* 38, 9533–9540.

33. Lee, K.-H. and Kuczera, K. (2003) Molecular dynamics simulation studies of cytochrome b_5 from outer mitochondrial and microsomal membrane. *Biopolymers,* 69, 260–269.

34. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comp. Chem. 4, 187-217.

35. MacKerell, A D ; Bashford, D; Bellott, M; Dunbrack, R L; Eva seck, J D; Field, M J; Fischer, S; Gao, J; Guo, H; Ha, S; Joseph McCarthy, D; Kuc nir, L; Kuczera, K;

Lau, F T K; Mattos, C; Michnick, S; Ngo, T; Nguyen, D T; Pro hom, B; Rehear, W E; Roux, B; Schlenkrich, M; Smith, J C; Stote, R; Straub, J; Wanabe, M; WiorkiewiczKuczera, J; Yin, D; Karplus, M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem., B 102, 3586-3617.

36. http://www.ks.uiuc.edu/

37. Feig, M., Karanicolas, J. and Brooks C.L, III, (2004), MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology
Journal of Molecular Graphics and Modeling, 22 (5), 377-95.

38. Feig, M., Im, W. and CL Brooks C.L, III, (2004), Implicit solvation based on generalized Born theory in different dielectric environments *J.Chem.Phys*. , 120 (2), 903-911 .

39. http://koehllab.genomecenter.ucdavis.edu/people/koehl/research/proshape/solvent-models

40. Altuve, A., Silchenko, S., Lee, K.-H., Kuczera, K., Terzyan, S., Zhang, X., Benson, D.R., and Rivera, M. (2001). Probing the differences between rat liver outer mitochondrial membrane cytochrome b_5 and microsomal cytochromes b_5. *Biochemistry*, 40, 9469–9483.

41. Konopka, K. and Waskell, L., (1988) Modification of trypsin-solubilized cytochrome b5, apocytochrome b5 and liposome-bound cytochrome b5 by diethyl pyrocarbonate. *Arch.Biochem.Biophys*., 261, 56-63.

42. Falzone, C.J., Wang, Y., Vu, B.C., Scott, N.L., Bhattacharya, S., and Lecomte, J.T.J. (2001) Structural and dynamic perturbations induced by heme binding in cytochrome b_5. *Biochemistry* ,40, 4879–4891.

43. Simeonov, .M. , Altuve, .A., Messiah, A.M, Wang, .A., Eastman, M.A, Benson, D.R and Rivera, M., (2005). Mitochodrial and Microsomal Ferric cytochrome b5 exhibit divergent conformational plasticity in the context of a common fold. *Biochemistry*, 44, 9308-9319.

44. Na Sun, Wang, A., Cowley, A.B., Altuve, A., Rivera, M., and Benson, D.R.,. (2005) Enhancing the stability of microsomal cytochrome b5 : a rational approach informed by comparative studies with outer Mitochondrial membrane isoform. *Protein Engineering Design and studies*, 18(12), 571-579.

45. Newbold, R.J. Hewson, R., Whitford, D., (1992) The thermal stability of the tryptic fragment of bovine microsomal cytochrome b5 and a variant containing six additional residues.*FEBS.Lett*. 314(3), 419-424.

46. Pfeil, W. (1993), Thermodynamics of apocytochrome b5: unfolding, *Prot.Sci.,* 2, 1497-1501

47. Manyusa, S., Gulnahar M., Whitford, D., (1999) Analysis of folding and unfolding reactions in cytochrome b5. *Biochemistry* 38, 14352-14362.

48. Rohl, C.A, Strauss, C.E.M., Chivian, D. and Baker , D. (2004) Modeling structurally variable regions in Homologous Proteins with Rosetta, Proteins, 55, 656-677

49. http://salilab.org/pdf/086_FiserDekker2000.pdf